The Astronomical Journal, 158:231 (57pp), 2019 December https://doi.org/10.3847/1538-3881/ab44a2
© 2019. The American Astronomical Society. All rights reserved.

The Data Analysis Pipeline for the SDSS-IV MaNGA IFU Galaxy Survey: Overview

Kyle B. Westfall[1], Michele Cappellari[2], Matthew A. Bershady[3,4], Kevin Bundy[1], Francesco Belfiore[1,5], Xihan Ji[6], David R. Law[7], Adam Schaefer[3], Shravan Shetty[3], Christy A. Tremonti[3], Renbin Yan[8], Brett H. Andrews[9], Joel R. Brownstein[10], Brian Cherinka[7,11], Lodovico Coccato[5], Niv Drory[12], Claudia Maraston[13], Taniya Parikh[13], José R. Sánchez-Gallego[14], Daniel Thomas[13], Anne-Marie Weijmans[15], Jorge Barrera-Ballesteros[11], Cheng Du[6], Daniel Goddard[13], Niu Li[6], Karen Masters[16], Héctor Javier Ibarra Medel[17], Sebastián F. Sánchez[17], Meng Yang[15], Zheng Zheng[18,19], and Shuang Zhou[6]

[1] University of California Observatories, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA; westfall@ucolick.org
[2] Sub-department of Astrophysics, Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH, UK
[3] Department of Astronomy, University of Wisconsin-Madison, 475N. Charter Street, Madison, WI 53703, USA
[4] South African Astronomical Observatory, P.O. Box 9, Observatory 7935, Cape Town, South Africa
[5] European Southern Observatory, Karl-Schwarzchild-Strasse 2, Garching bei München, D-85748, Germany
[6] Tsinghua Center of Astrophysics & Department of Physics, Tsinghua University, Beijing 100084, People's Republic of China
[7] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[8] Department of Physics and Astronomy, University of Kentucky, 505 Rose Street, Lexington, KY 40506, USA
[9] University of Pittsburgh, PITT PACC, Department of Physics and Astronomy, Pittsburgh, PA 15260, USA
[10] University of Utah, Department of Physics and Astronomy, 115 S. 1400 E., Salt Lake City, UT 84112, USA
[11] Center for Astrophysical Sciences, Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA
[12] McDonald Observatory, The University of Texas at Austin, 1 University Station, Austin, TX 78712, USA
[13] Institute of Cosmology & Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth, PO1 3FX, UK
[14] Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195, USA
[15] School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews, KY16 9SS, UK
[16] Department of Physics and Astronomy, Haverford College, 370 Lancaster Ave, Haverford, PA 19041, USA
[17] Instituto de Astronomia, Universidad Nacional Autónoma de México, A.P. 70-264, 04510, Mexico, D.F., México
[18] National Astronomical Observatories of China, Chinese Academy of Sciences, 20A Datun Road, Chaoyang District, Beijing 100012, People's Republic of China
[19] CAS Key Laboratory of FAST, NAOC, Chinese Academy of Sciences, People's Republic of China
Received 2019 January 3; revised 2019 August 31; accepted 2019 September 13; published 2019 November 18

Abstract

The Mapping Nearby Galaxies at Apache Point Observatory (MaNGA) survey is currently acquiring integral-field spectroscopy for the largest sample of galaxies to date. By 2020, the MaNGA Survey—which is one of three core programs in the fourth-generation Sloan Digital Sky Survey (SDSS-IV)—will have observed a statistically representative sample of 10^4 galaxies in the local universe ($z \lesssim 0.15$). In addition to a robust data-reduction pipeline (DRP), MaNGA has developed a data-analysis pipeline (DAP) that provides higher-level data products. To accompany the first public release of its code base and data products, we provide an overview of the MaNGA DAP, including its software design, workflow, measurement procedures and algorithms, performance, and output data model. In conjunction with our companion paper (Belfiore et al.), we also assess the DAP output provided for 4718 observations of 4648 unique galaxies in the recent SDSS Data Release 15 (DR15). These analysis products focus on measurements that are close to the data and require minimal model-based assumptions. Namely, we provide stellar kinematics (velocity and velocity dispersion), emission-line properties (kinematics, fluxes, and equivalent widths), and spectral indices (e.g., D4000 and the Lick indices). We find that the DAP provides robust measurements and errors for the vast majority (>99%) of analyzed spectra. We summarize assessments of the precision and accuracy of our measurements as a function of signal-to-noise. We also provide specific guidance to users regarding the limitations of the data. The MaNGA DAP software is publicly available and we encourage community involvement in its development.

Key words: methods: data analysis – techniques: imaging spectroscopy – surveys – galaxies: general – galaxies: fundamental parameters

1. Introduction

Publicly available datasets in accessible formats play an increasingly important role in astronomy; for example, they broaden access and enable analyses that combine observations across wavelengths and telescopes. The original Sloan Digital Sky Survey (SDSS, York et al. 2000) initiated an ongoing commitment to publicly release raw and reduced data that continues through the current generation, SDSS-IV (Blanton et al. 2017).

With the introduction of the MaNGA Survey (Mapping Nearby Galaxies at Apache Point Observatory, Bundy et al. 2015), SDSS-IV data releases have included higher-dimensional sets of data cubes whose production requires a sophisticated Data Reduction Pipeline (Law et al. 2016, DRP). Over its six-year duration, ending 2020 July, MaNGA aims to provide spatially resolved spectroscopy for 10,000 nearby galaxies selected with $\mathcal{M}_* \gtrsim 10^9 \mathcal{M}_\odot$ and $\langle z \rangle \sim 0.03$ (Wake et al. 2017). MaNGA uses specially designed fiber bundles (Drory et al. 2015) that feed the BOSS spectrographs (Smee et al. 2013) on the 2.5 m Sloan Telescope (Gunn et al. 2006). Spanning 0.36–1.0 μm at a resolution of $R \sim 2000$ and with excellent flux calibration (Yan et al. 2016b), MaNGA executes approximately three-hour long dithered integrations (Law et al. 2015) to reach signal-to-noise ratio (S/N) requirements for galaxies observed to

approximately 1.5 effective (half-light) radii, R_e, (two-thirds of the sample) and 2.5 R_e (one-third of the sample) (Yan et al. 2016a).

Reduced MaNGA data, including reconstructed data cubes, are produced by the automated MaNGA DRP and have been made publicly available since the thirteenth SDSS data release (DR13; Albareti et al. 2017). Inspired by previous "value-added catalogs" (VACs) based on SDSS data, with the MPA-JHU[20] catalog for SDSS-I/II being a prime example, members of the MaNGA team have provided publicly available VACs in previous data releases.[21]

The complexity and richness of spatially resolved datasets, such as those produced by MaNGA, as well as the desire to take on common analysis tasks that would otherwise be duplicated by large numbers of users has motivated SDSS-IV to invest in a "project-led" MaNGA Data Analysis Pipeline (DAP). By providing a uniform set of commonly desired analysis products, the DAP also enables rapid and interactive delivery methods for these data products, helping scientists to quickly design samples of interest and make discoveries. It was also appreciated that while researchers would likely perform custom measurements of primary interest to their science, the ability to combine these with readily available DAP measurements that might otherwise be outside their expertise would open up new opportunities. In general, a broad-based and robust DAP makes MaNGA data more science-ready. This strategy aligns with other integral-field unit (IFU) surveys that have released high-level data products, including ATLAS3D[22] (Cappellari et al. 2011), CALIFA[23] (Sánchez et al. 2012), and SAMI[24] (Croom et al. 2012).

The MaNGA DAP has been under development since 2014, evolving through several versions and a transition from an original IDL to Python implementation. Although the DAP has primarily been used as a survey-level pipeline for providing analysis products to the SDSS collaboration, our development strategy has also emphasized flexibility to prospective users. The low-level, core algorithms have been constructed in a way that is largely independent of their specific use with the MaNGA data, which allows a user to write new Python scripts around DAP functions or classes for analysis of more varied datasets. Additionally, the high-level interface is written such that the detailed execution of many of the internal algorithms can be modified using a set of configuration files, which allows a user to tailor how the DAP analyses MaNGA data to better suit their scientific needs. Although much of the high-level functionality assumes one is working with MaNGA data, it is possible to apply the DAP to data from different instruments with modest modification.

The MaNGA science teams have published studies based on DAP output throughout its development, using internal data releases to the SDSS collaboration that we term MaNGA Product Launches (MPLs). The DAP source code continues to evolve as we improve its fidelity and expand its functionality. We encourage community involvement in these efforts via our public repository on GitHub.[25] The GitHub repository also contains some example scripts that use low-level DAP functions to fit a single spectrum, as well as the higher-level modules (Section 4) that can be used to fit a single data cube. Additional example scripts will be provided as development continues and we encourage others to submit their own scripts via a GitHub pull request.

This paper describes the first public release of DAP products, which is part of SDSS Data Release 15 (DR15), and includes both the code (version 2.2.1) and output data products. The public DAP output is available for download from the SDSS website[26] and via Marvin[27] (Cherinka et al. 2019). Marvin is a Python package that provides an interactive web-based interface to both MaNGA data cubes and DAP quantities (Marvin-Web), as well an API (Marvin-API) and a user-level toolset (Marvin-tools) that enables seamless remote and/or local access to and interaction with these MaNGA data in any Python-based analysis workflow.

The primary output of the DAP includes stellar kinematics, fluxes and kinematics of emission lines, and continuum spectral indices. In deriving these measurements, the DAP makes heavy use of pPXF (Cappellari & Emsellem 2004; Cappellari 2017) as a workhorse spectral-fitting routine. Our philosophy has been to focus on measurements that are made directly on the MaNGA spectra and that do not require significant model-based assumptions. For example, the DAP fits stellar template mixes plus a polynomial component to the stellar continuum. Although this provides an excellent representation of the data and derived stellar absorption-line kinematics, it is not necessarily appropriate for accurate stellar-population properties. For maps of estimated stellar age, metallicity, star formation histories, and other model-derived data products, the Firefly (Goddard et al. 2017) and Pipe3D (Sánchez et al. 2016a, 2016b) VACs[28] are valuable resources. In particular, Pipe3D products also include alternative measurements of kinematics and emission lines.

In this contribution, we present an overview of the MaNGA DAP, its algorithmic structure, and its output data products. We begin in Section 2 with a "quick-start" guide, which provides a more detailed introduction to the DAP output and nomenclature by way of examples. Along with many other resources cited therein, we expect Section 2 to be a useful road map to the DAP products and the rest of our paper. In Section 3, we describe the input data required by the DAP, including a general summary of MaNGA spectroscopy. Section 4 provides an overview of the DAP workflow through its six main analysis modules. After first describing the spectral templates we use in DR15 and the method used to generate them in Section 5, we dedicate the following five sections to the detailed description of each of the DAP analysis modules.[29]

Section 6 describes our spatial-binning approach and the importance of including spatial covariance in these calculations. Section 7 goes into particular depth with regard to the assessments of the stellar kinematics. In particular, we present results from several input/output simulations, as well as a

[20] https://www.sdss.org/dr15/data_access/value-added-catalogs/mpa-jhu-stellar-masses
[21] For example: https://www.sdss.org/dr14/data_access/value-added-catalogs/.
[22] http://purl.org/atlas3d
[23] http://califa.caha.es
[24] https://sami-survey.org
[25] https://github.com/sdss/mangadap
[26] https://www.sdss.org/dr15/data_access
[27] https://dr15.sdss.org/marvin
[28] https://www.sdss.org/dr15/data_access/value-added-catalogs/
[29] In these sections, we focus on the specific way that the DAP has been executed for the data provided as part of DR15, not on an exhaustive description of what the DAP *can* do. More exhaustive and evolving documentation of the code is provided as part of the code distribution and is hosted at https://sdss-mangadap.readthedocs.io/en/latest/.

statistical comparison of the results for MaNGA galaxies with multiple observations. These repeat observations allow us to provide a detailed assessment of the errors reported by the DAP. We use Section 8 to introduce our bandpass-integral formalism that is used when measuring both non-parameteric emission-line fluxes and spectral indices. Section 9 describes our emission-line modeling algorithm, with detailed assessments of the module and its usage in DR15 provided by our companion paper, Belfiore et al. (2019). Finally, Section 10 describes our quantification of continuum features using spectral indices, and we assess the accuracy and precision of our measurements using a similar approach used for the stellar kinematics. Detailed quality assessments of the two full-spectrum-fitting modules are provided in Section 7—stellar kinematics—and in Belfiore et al. (2019)—emission-line modeling.

We comment on the overall performance of the DAP for DR15 in Section 11. In particular, we note specific regimes where we find the DAP currently requires further development. Section 12 provides a detailed discussion of the DAP output products and important aspects of these products that users should keep in mind. Finally, we provide some brief conclusions in Section 13. Appendices A, B, and C provide, respectively, the DAP procedure used to match the spectral resolution of two spectra, an assessment of how instrumental resolution errors propagate to errors in the stellar velocity dispersion measurements, and tables that describe the DAP output data models.

Unless stated otherwise, throughout this paper: (1) we adopt a ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 100h$ km s^{-1} Mpc^{-1}; (2) all wavelengths are provided in vacuum; and (3) all flux densities have units of 10^{-17} erg cm^{-2} s^{-1} Å^{-1} spaxel^{-1}.

2. DAP Quick-start Guide

We begin with a "quick-start guide" to the MaNGA DAP, jumping right into example output products that highlight key aspects of the DAP measurements. To be clear from the beginning, *all* of the DAP input and output files that we discuss here and throughout our paper are included as part of DR15. The narrative of this section aims to help the reader to quickly get a sense of those products and to navigate their way to the subsequent sections of our paper that are most relevant to their goals. Although some general guidance can be sufficiently provided here, other more nuanced advice requires the backdrop of our assessments of the DAP data, which are performed throughout our paper. In Section 2.1, we highlight specific sections where the reader can go for that advice, as well as other documentation provided as part of SDSS DR15. Finally, Section 2.2 provides a list of known issues with the DAP data provided with DR15, which users should take into account.

In Figure 1, we show the SDSS *gri* composite image for the galaxy targeted by IFU 12703 on plate 8439 and a sampling of the quantities produced by the DAP for this data cube.[30] From top to bottom and left to right, these images roughly follow the order of the DAP workflow (Section 4), from assessments of the *g*-band signal-to-noise ratio (S/N_g per spectral pixel; panel (b); Section 6); to the measurements of the stellar kinematics (panels (c) and (d); Section 7); to the emission-line modeling that produces fluxes, equivalent-widths (EWs), and kinematics (panels (e) through (l); Section 9; Belfiore et al. 2019); and finally to the spectral-index measurements (panels (m) through (r); Section 10).

The images, or maps, plotted in Figure 1 are provided by the primary DAP output file, which is a multi-extension fits file called the MAPS file (Section 12.1). The MAPS files provide each DAP measurement in a two-dimensional image format, or map, that exactly matches the spatial dimensions of the DRP data cubes. Where appropriate, each mapped quantity has associated inverse-variance and quality-assessment measurements. Our quality assessments are provided by a set of bitmasks, which are defined in Appendix C, and their use is a critical aspect of any workflow incorporating the DAP output data. In Figure 1, the empty regions are masked either because they are outside the hexagonal footprint of MaNGA's dithered field of view (FOV) or because they do not meet the DAP quality-assurance criteria (Section 6.1). For Figure 1 specifically, we also mask regions in the maps of the stellar velocity dispersion (σ_*), Hα velocity dispersion ($\sigma_{H\alpha}$), and spectral indices at low S/N_g and low flux; see the figure caption.[31] A user-customized set of DAP maps are easily displayed for a specific galaxy in Marvin, both via the web interface[32] and its Python toolset.[33]

The DAP results for 8439-12703 (the data cube for galaxy 1-605884) were chosen at random for Figure 1 and are representative of the DR15 results as a whole.[34] Using these data as an example, it is important to note that the spatial pixel, or spaxel, size ($0\farcs5 \times 0\farcs5$) is significantly smaller than the full-width at half maximum (FWHM) of the on-sky point-spread function ($2\farcs5$ diameter) shown as a gray circle in the bottom-left corner of each panel. This spaxel sampling was chosen as a compromise between the covariance introduced in the data cubes by our reconstruction approach and the spatial scale needed to properly sample the dithered fiber observations (see Liu et al. 2019). Even so, the subsampling of the fiber beam leads to significant covariance between adjacent spaxels (Section 6.2; Law et al. 2016) and a few noteworthy implications. First, accounting for this covariance is critical to accurately meeting a target S/N threshold when spatially binning the data cubes (Section 6.3). Second, spatial variations in the mapped measurements driven by random errors in the observed spectra should be smooth on scales of roughly 5 × 5 spaxels. That is, significant spaxel-to-spaxel variations due to a random sampling are highly improbable given the well-defined correlation matrix of the data cube. Instead, a useful rule of thumb to keep in mind when inspecting DAP images is that significant spaxel-to-spaxel variations are driven by systematic error,[35] not astrophysical structure. Meanwhile, structure on scales similar to the beam size could be astrophysical—such as

[30] The Python code used to produce this plot and many others in this paper can be found at https://github.com/sdss/mangadap/tree/master/docs/papers/Overview/scripts.

[31] Masking in the DAP is largely limited to indicating numerical or computational issues occurring during the course of the analysis. Flagging of any given measurement based on its *expected* quality is more limited due to the difficulty of defining criteria that are generally robust and not overly conservative.

[32] https://dr15.sdss.org/marvin

[33] https://github.com/sdss/marvin

[34] Because the same galaxy may be observed more than once (see Table 1), we generally refer to a specific data cube using its PLATEIFU designation throughout this paper (e.g., 8439-12703 in this case), as opposed to the unique MaNGA ID associated with each survey target.

[35] The systematic error involved may only yield an increased stochasticity in the measurements that average out over many spaxels, and does not necessarily imply a systematic shift of the posterior distribution away from the true value.

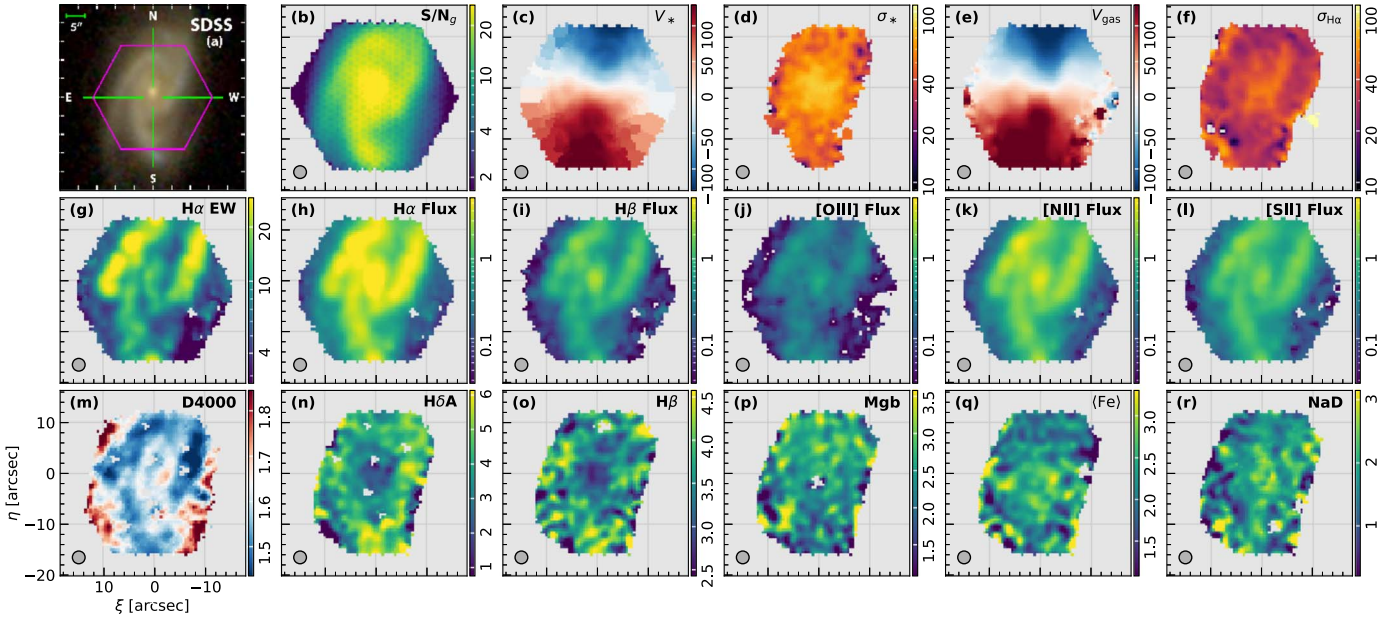

Figure 1. A subset of the DAP-derived quantities for data cube 8439-12703, the observation of MaNGA galaxy 1-605884 ($z = 0.025$; $\log(\mathcal{M}_*/h^{-2}\mathcal{M}_\odot) = 10.2$; $R_e = 11.''8 \sim 4.2\ h^{-1}$ kpc; see https://sas.sdss.org/marvin/galaxy/8439-12703/ for more information about this galaxy and its properties in the larger context of the MaNGA sample). From top-to-bottom and left-to-right: (a) the SDSS *gri* composite image with the nominal size of the IFU outlined in purple; (b) the *g*-band S/N per channel, S/N$_g$; (c) the stellar line-of-sight (LOS) velocity, V_*; (d) the stellar velocity dispersion, σ_*; (e) the ionized-gas LOS velocity, V_{gas}; (f) the velocity dispersion of the Hα emission line, $\sigma_{H\alpha}$; (g) the equivalent width (EW) of the Hα emission line; (h) the flux of the Hα emission line; (i) the flux of the Hβ emission line; (j) the total flux in the [O III]λ4959,5007 emission lines; (k) the total flux in the [N II]λ6548,6583 emission lines; (l) the total flux in the [S II]λ6716,6730 emission lines; (m) the D4000 spectral index; (n) the HδA spectral index; (o) the Mgb spectral index; (q) the average of the Fe5270 and Fe5335 spectral indices, $\langle$Fe$\rangle$; and (r) the NaD spectral index. The gray circle in the bottom-left corner of panels (b) through (r) is the nominal FWHM of MaNGA's spatial resolution element (beam size; $2.''5$). Bins or spaxels with S/N$_g < 10$ are masked in panel (d) (σ_*); S/N$_g < 3$ are masked in panel (m) (D4000); S/N$_g < 5$ are masked for the plotted absorption-line indices (panels (n) through (r)); and Hα fluxes less than 2.5×10^{-18} erg s^{-1} cm^{-2} are masked in panel (f) ($\sigma_{H\alpha}$).

the increased Hα EW along the spiral arms of galaxy 1-605884—or driven by noise in the fiber observations—such as is likely the cause of the strong variations in the gas velocity field toward the IFU periphery or the high-frequency modulations of its spectral-index maps.

The maps shown in Figure 1 are the result of a "hybrid" binning approach (see the introduction to Section 9 and the algorithm description in Section 9.2). In this approach, the stellar kinematics are measured for spatially binned spectra that meet a minimum of S/N$_g \gtrsim 10$ (Section 6.3) and the emission-line and spectral-index measurements are performed for individual spaxels. We expect these results to be preferable for the majority of users, providing the benefits of both unbiased stellar kinematics and unbinned emission-line maps. The selection of these products is made via the DAPTYPE, as we define below.

A core design principle of the DAP has been to abstract and modularize the analysis steps to maintain flexibility. The specific combination of the settings used for all of the analysis steps (e.g., the specific binning algorithm and the templates used to measure the stellar kinematics) adopted for the survey-level execution of the DAP is used to construct a unique keyword called the DAPTYPE; see Section 4 and Figure 3 for more detail. For DR15, two unique approaches to the analysis were performed, meaning that each galaxy data cube is analyzed twice and users must choose which set of products to use (see Section 12).[36] The fundamental difference between these two DAPTYPEs is whether the output is based on the hybrid-binning approach—the output we recommend users start with—or if *all* analyses have been performed using spectra Voronoi-binned to a target S/N$_g \gtrsim 10$.

To minimize the complexity of the output DAP data model, measurements made on binned spectra are mapped to each spaxel in the bin. This can be seen in the maps of the stellar kinematics shown in Figure 1, where regions of constant stellar velocity are the visual result of all the relevant spaxels being binned into a single spectrum during the fitting process. Alternatively, the remaining maps all show quantities varying spaxel-to-spaxel. If the maps in Figure 1 were instead from the analysis that only used the Voronoi-binned spectra (not the hybrid approach), then *all* of the maps would show identical regions of constant values. Although they are convenient for visual inspection and for the simplicity of the data model, it is important to identify and select only the unique measurements for detailed analysis (see other data-model-specific advice in Section 12).

Inevitably, visual inspection of the DAP maps will lead one to find features that are physically counter-intuitive and/or erroneous; we discuss some of these cases in Section 11.2.2. When in doubt about features seen in the DAP maps, there is no substitute for directly inspecting the MaNGA spectral data (Section 3; Law et al. 2016) and the associated DAP model spectra. The latter are provided by the second main DAP output file, which is a multi-extension fits file called the model LOGCUBE file (Section 12.2). This file contains all the model spectra fit to the observations, or provides the information needed to reconstruct the models (see point 2 in Section 12.2). Our Marvin software package and web-based interface

[36] For reference, the execution time of the DAP strongly depends on the number of spectra being analyzed, such that larger IFU bundles require more time. For a single core of our Utah cluster (comparable to a single core on a laptop), the median execution time to complete both analysis approaches on a single data cube for DR15 was 2.0, 3.1, 4.6, 6.6, and 8.7 hr for the 19-, 37-, 61-, 91-, and 127-fiber data cubes, respectively. However, some of the 127-fiber data cubes required up to 25 hr to complete.

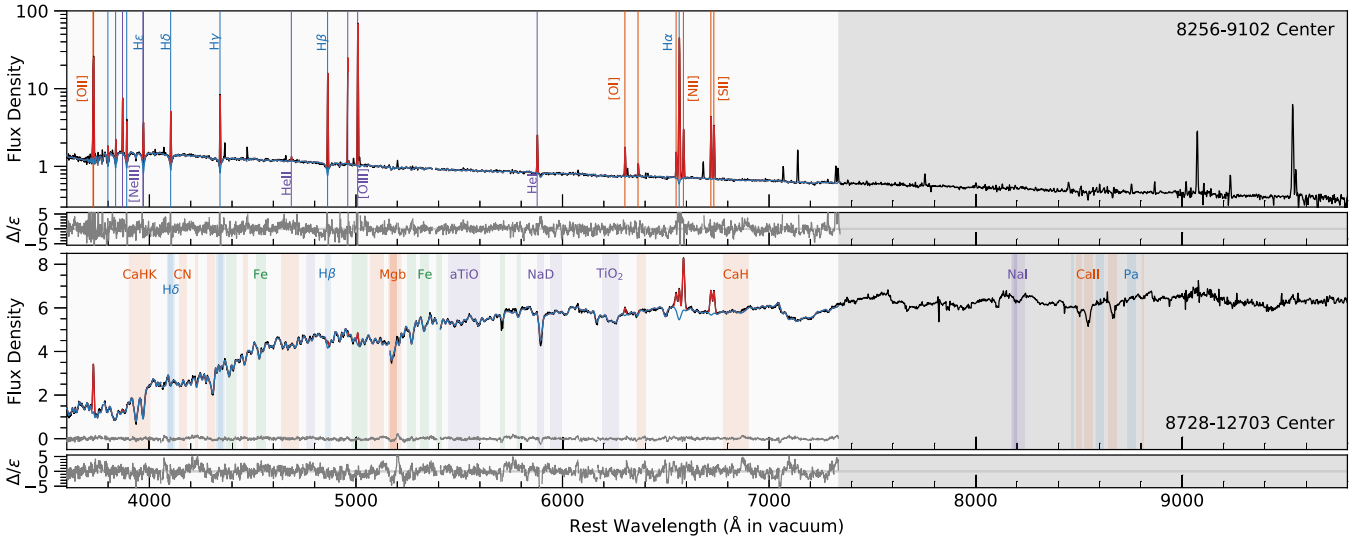

Figure 2. The central spaxel of data cubes `8256-9102` (top) and `8728-12703` (bottom) and the best-fitting `DAP` model spectra. Flux densities are plotted in units of 10^{-17} erg s^{-1} cm^{-2} Å^{-1} spaxel^{-1}. In both panels, the observed spectrum is shown in black, the best-fitting model spectrum (stellar continuum plus emission lines) is shown in red, and the stellar-continuum-only model is shown in blue. The fit residuals are shown directly for `8728-12703`, and also shown in separate panels for both spectra after normalizing by the spectral errors, Δ/ϵ. The gray region at $\lambda \gtrsim 7340$ Å is not included in any full-spectrum fit as the stellar spectral templates used for DR15 have no coverage at these wavelengths. The 22 emission lines fit in DR15 (Table 3) are labeled and marked in the top panel by vertical lines colored according to the groups identified in Section 5.3 of Belfiore et al. (2019). The primary passbands of the 43 absorption-line indices measured in DR15 (Table 4) are shown against the `8728-12703` spectrum: Hydrogen bands are marked in blue; C, N, Ca, or Mg bands are in orange; Fe bands are in green; and Na or TiO bands are in purple. A few bands are labeled according to their name or the element present in their name.

(Cherinka et al. 2019) provide particularly useful tools for this kind of data inspection.

Figure 2 shows two high-S/N$_g$ spectra and the best-fitting `DAP` model spectra. Note that the full model (stellar continuum and emission lines) is shown in red with the stellar-continuum-only model overlaid in blue; the two models only differ by the presence of the emission lines such that they both appear blue in most of the figure. The model spectra are the result of the two full-spectrum fits in the `DAP`, both of which use `pPXF`. The first fit masks the emission lines and determines the best fit to the stellar continuum to measure the stellar kinematics (Section 7). The second fit simultaneously optimizes the stellar continuum and emission lines, while keeping the stellar kinematics fixed to the result from the first fit (Section 9). The stellar continuum is handled differently between these two fits and yields slightly different results; the continuum shown in Figure 2 is the result of the combined fit of the second full-spectrum fit (see Belfiore et al. 2019, Figure 2).

Note that the wavelength range fit by the `DAP` full-spectrum-fitting modules is limited to 0.36–0.74 μm for DR15 because of the spectral range of the templates used (Section 5). This has two primary effects. First, it limits the spectral range over which we can fit emission lines. Most notably this excludes modeling of the near-infrared [S III] lines. For this purpose, in particular, we aim to soon take advantage of our in-house stellar library, MaStar[37] (Yan et al. 2019), so that our continuum models are fit over MaNGA's full spectral range. Second, while we can measure spectral indices at all wavelengths, we can only calculate the velocity-dispersion corrections (Section 10.1) for those measurements in regions with valid model fits. This means any spectral index provided in DR15 with a main passband centered at $\lambda > 0.74$ μm does not include a velocity-dispersion correction, which can be critical when, e.g., analyzing absorption-line strengths as a function of galaxy mass.

The two galaxy spectra in Figure 2 were selected to illustrate the features fit by the `DAP`. The central spaxel of data cube `8256-9102` for the star-forming galaxy 1-255959 has extremely bright nebular emission with nearly all 22 emission lines measured in DR15 (Sections 9 and 8; Table 3) identifiable by eye. Indeed, many more emission lines are visible that are not currently fitted by the `DAP`, largely from H and He recombination and N, O, S, and Ar forbidden transitions. We expect to add to the list of lines included in the fit in future releases of the `DAP`, both by extending the spectral range of the continuum models (see Belfiore et al. 2019) and controlling for any problems caused by attempting to fit what are generally much weaker lines. The central spaxel of data cube `8728-12703` of the early-type galaxy 1-51949 has relatively weak emission features but exhibits many of the absorption features measured by the 46 spectral indices provided in DR15 (Section 10 and Table 4).

2.1. Usage Guidance

Anyone planning to use the `DAP` data is strongly encouraged to read Section 12, where we describe the main output files provided by the `DAP` and we highlight a number of aspects of the data important to their use. Some of these are simple practicalities of the data model, but others are critical to the proper interpretation of the data. Basic introductions and usage advice for the `DAP` data products are also included in the DR15 paper (Aguado et al. 2019) and data-release website, https://www.sdss.org/dr15/. From the latter, note the following in particular: an introduction to working with MaNGA data (https://www.sdss.org/dr15/manga/getting-started/) and some of the intricacies involved (https://www.sdss.org/dr15/manga/manga-data/working-with-manga-data/), worked tutorials (https://www.sdss.org/dr15/manga/manga-tutorials/, a list of

[37] https://www.sdss.org/surveys/mastar/

known problems and caveats (https://www.sdss.org/dr15/manga/manga-caveats/), and the general SDSS helpdesk (https://www.sdss.org/dr15/help/).

In terms of its general success in fitting MaNGA spectra, Section 11 provides a useful reference if one encounters a missing DAP product or a counter-intuitive measurement. In particular, Section 11.2.2 provides a list of regimes where fit-quality metrics have been used to identify aberrantly poor spectral fits produced by the DAP.

Guidance for each of the three primary DAP product groups (stellar kinematics, emission-line properties, and continuum spectral indices) are provided in, respectively, Section 7, Belfiore et al. (2019), and Section 10. For the stellar kinematics, we particularly recommend Section 7.7, which provides guidance for how to use our stellar velocity dispersion measurements. For the emission-line measurements, we recommend that users read Sections 6 and 7 of Belfiore et al. (2019), at least, and then follow-up with other Sections of their paper as relevant. Finally, for the spectral indices, we recommend users read the summary of the assessments that we have performed herein, which is provided in Section 10.3.4.

2.2. Known Issues in DR15 DAP Products

For general reference, we provide a list of known issues with the DR15 version of the DAP software and the DAP-derived data products. Where relevant, we provide references to subsequent sections of our paper with more information; see also https://www.sdss.org/dr15/manga/manga-caveats/. More up-to-date information and documentation of source code changes are included in the source-code distribution, see https://github.com/sdss/mangadap/blob/master/CHANGES.md.

1. Detailed assessments of the uncertainties provided for the non-parametric emission-line measurements have not been performed, meaning that their accuracy is not well characterized. We expect that they are of similar quality to the spectral-index uncertainties (Section 10.3.2); however, they should be treated with caution.
2. Some Milky Way foreground stars that fall within MaNGA galaxy bundles have not been properly masked. More generally, the DAP does not correctly handle the presence of multiple objects in the IFU bundle FOV (Section 11.2.2).
3. The wings of particularly strong or broad emission lines will not have been properly masked during the stellar-continuum fits used to measure the stellar kinematics (Section 7.1.2).
4. The MASK extension in the model LOGCUBE files cannot be used to reconstruct the exact mask resulting from the stellar-kinematics module.
5. The χ^2_ν measurements reported for the stellar-kinematics module are not correct—they do not exclude pixels that were rejected during the fit iterations.
6. The highest order Balmer line fit by the DAP is Hθ, even though many galaxies show higher-order lines (Figure 2).
7. Measurements for the Hζ Balmer line are unreliable given its blending with the nearby He I line (as reported by Belfiore et al. 2019).
8. The velocity-dispersion measurements for the [O II] line are improperly masked, which is also reported by Belfiore et al. (2019).
9. The DAPall file (Section 12.3) reports spectral indices within 1 R_e that have *not* been corrected for the observed stellar velocity dispersion (Section 10.1).
10. Velocity-dispersion corrections for index measurements will include velocity effects because all measurements are done using the single bulk redshift to offset the band definitions (Section 10).

3. DAP Inputs

3.1. MaNGA Spectroscopy

Drory et al. (2015) provide a detailed description of the MaNGA fiber-feed system, which is composed of 17 IFUs: two 19-fiber IFUs, four 37-fiber IFUs, four 61-fiber IFUs, two 91-fiber IFUs, and five 127-fiber IFUs. The plate scale of the 2.5 m Sloan telescope yields an on-sky fiber diameter of 2″. The combination of the seeing conditions at Apache Point Observatory (APO), the dithering pattern of the MaNGA observational strategy (Law et al. 2015), and the method used to construct the data cubes typically provides a spatial point-spread function (PSF) with a FWHM of ∼2.″5 (Law et al. 2016). All of the IFUs have fibers that are packed in a hexagonal, regular grid with a FOV directly related to the number of fibers. Including the fiber cladding, the nominal FOV diameters are 12″, 17″, 22″, 27″, and 32″ for the 19-, 37-, 61-, 91-, and 127-fiber IFUs, respectively.

The MaNGA fiber-feed systems are coupled to the SDSS-III/BOSS spectrographs (Smee et al. 2013), a pair of spectrographs with "blue" and "red" cameras that receive, respectively, reflected ($\lambda \lesssim 0.63\,\mu$m) and transmitted ($\lambda \gtrsim 0.59\,\mu$m) light from a dichroic beamsplitter. The full spectral range obtained for each fiber spectrum is $0.36\,\mu\text{m} \lesssim \lambda \lesssim 1.03\,\mu$m after combining the data from both cameras. Each arm of each spectrograph uses a volume-phase holographic grism yielding spectral resolutions of $R_\lambda = \lambda/\Delta\lambda \approx 2000$ at $\lambda = 0.55\,\mu$m for the two blue cameras and $R_\lambda \approx 2500$ at $\lambda = 0.9\,\mu$m for the two red cameras (see Yan et al. 2016a, Figure 20).

Following the observational strategy outlined by Law et al. (2015), each MaNGA plate is observed using a three-point dither pattern to fill the IFU interstitial regions and optimize the uniformity of the FOV sampling for all 17 targeted galaxies on a plate. Depending on the observing conditions, 2–3 hr of total observing time is required to reach the survey-level constraints on the S/N, as defined by Yan et al. (2016a).

These data are reduced by the MaNGA DRP, an IDL-based software package, yielding wavelength-, flux-, and astrometrically calibrated spectra. The reduction procedures are similar to those used by the SDSS-III/BOSS pipeline,[38] but with significant adjustments as required by the MaNGA observations. The DRP is described in detail by Law et al. (2016), the spectrophotometric calibration technique is described by Yan et al. (2016b), and relevant updates to these procedures for DR15 are discussed by Aguado et al. (2019).

For each fiber, the spectra from the "blue" and "red" cameras are combined into a single spectrum and resampled to a common wavelength grid. Spectra are produced with both linear and log-linear wavelength sampling. All spectra for a given PLATEIFU designation are included in a single file as a set of row-stacked spectra (RSS) and as a uniformly sampled

[38] https://www.sdss.org/dr15/spectro/pipeline/

data cube (CUBE). The MaNGA data cubes are constructed by regridding the flux in each wavelength channel to an on-sky pixel (spaxel) sampling of 0.″5 on a side following the method of Shepard (1968); see Law et al. (2016, Section 9) for details (cf. Sánchez et al. 2012; Liu et al. 2019). The interpolating kernel is a two-dimensional Gaussian with a standard deviation of 0.″7 and a truncation radius of 1.″6. Although this interpolation process leads to significant covariance between the spaxels in a given wavelength channel (Law et al. 2016, Section 9.3), the current DAP release is primarily focused on working with these resampled data cubes. In addition, the DAP currently only analyzes the spectra that are sampled with a log-linear step in wavelength of $\Delta \log \lambda = 10^{-4}$ (corresponding to a velocity scale of $\Delta V = 69$ km s^{-1}).

3.2. Photometric Metadata

For convenience, the DAP uses measurements of the ellipticity ($\epsilon = 1 - b/a$) and position angle (ϕ_0) of the *r*-band surface-brightness distribution to calculate the semimajor-axis elliptical polar coordinates, R and θ, where the radius is provided in arcseconds, as well as in units of the effective (half-light) radius, R_e. In the limit of a tilted thin disk, these are the in-plane disk radius and azimuth. Except for some targets from MaNGA's ancillary programs, the photometric data are taken from the parent targeting catalog described by Wake et al. (2017, Section 2). This catalog is an extension of the NASA-Sloan Atlas (NSA)[39] toward higher redshift ($z \lesssim 0.15$) and includes an elliptical-Petrosian analysis of the surface-brightness distributions. Despite this difference with respect to the NSA catalog provided by the catalog website, we hereafter simply refer to this extended catalog as the NSA. The primary advantage of the NSA is its reprocessing of the SDSS imaging data to improve the sky-background subtraction and to limit the "shredding" of nearby galaxies into multiple sources (Blanton et al. 2011).

4. Workflow

At the survey level, the DAP is executed once per DRP data cube (PLATEIFU). The DAP will attempt to analyze any data cube produced by the DRP, as long as it is an observation of a galaxy target that has an initial estimate of its redshift. Specifically, we only analyze observations selected from the DRPall file (Law et al. 2016) for galaxies in either the main MaNGA survey or its ancillary programs[40] (respectively, either mngtarg1 or mngtarg3 are non-zero) and with a redshift of $cz > -500$ km s^{-1}. The restriction on the redshift is required because of the ± 2000 km s^{-1} limits we impose in the fit of the kinematics for each observation (Sections 7 and 9); we allow galaxies with a small blueshift so that the DAP will analyze a few observations of local targets from ancillary programs. Importantly, the DAP will analyze data cubes that the DRP has marked as critical failures (DRPQUAL is CRITICAL in the DRPall file); the appropriate flag is propagated to the global DAP quality bit (DAPQUAL; see Appendix C). As noted by Law et al. (2016, Section B.4), data cubes marked with CRITICAL quality flags should be used with caution or simply omitted from use. However, the approach to flagging reductions as CRITICAL is purposely conservative, meaning that some of these reductions may yet be valid. Therefore, we simply include them in our DAP analysis but caution users similarly concerning their use. The most conservative approach is for users to ignore data marked as CRITICAL by MaNGA quality bits. In total, the DAP has analyzed 4731 observations for DR15.

The primary DRP-produced output passed to the DAP for analysis are the MaNGA data cubes that are sampled logarithmically in wavelength (i.e., the DRP LOGCUBE files).[41] The DAP uses two additional text files to set its execution procedures: the first provides the photometric and redshift data for each target, which are most often drawn from the NSA (Section 3.2); and the second defines a set of "analysis plans" that are executed in sequence. We refer to each analysis plan as the DAPTYPE of a given output dataset. An analysis plan is composed of a set of keywords that select preset configurations of the low-level parameters that dictate the behavior of each DAP module, with one keyword per primary module. The six primary modules of the DAP have the following analysis goals: (1) perform basic data-quality assessments and calculations using the DRP-produced data (Sections 6.1 and 6.2), (2) spatially bin the DRP data cube (Sections 6.3 and 6.4), (3) measure the stellar kinematics (Section 7), (4) use bandpass integrals to compute non-parametric moments of the emission lines (Section 9.7; Belfiore et al. 2019), (5) fit parametric models to the emission lines (Section 9), and (6) use bandpass integrals to compute a set of absorption-line and bandhead indices (Section 10).

Figure 3 provides a schematic of the DAP workflow through its six modules. The modules are executed in series, from top to bottom in the figure, with each module often depending on the results of all the preceding modules. All modules are executed once per DAPTYPE, with the exception of the emission-line moment calculation (the fourth module), which is run once before and once after the emission-line model fitting (see Sections 9 and 8 for more detail). The analysis objectives of each module are also listed in the figure.

The results of each module are saved in its "reference file." These reference files: (1) allow the DAP to reuse (as opposed to recompute) analysis results common to multiple analysis plans (e.g., using the same data-quality assessments from the first module with different binning schemes), (2) allow the DAP to effectively restart at the appropriate module in case of a failure, and (3) provide access to a more extensive set of data beyond what is currently propagated to the two main output files, the MAPS and model LOGCUBE files. The reference files are released as part of DR15 with their data model documented at the DR15 website;[42] however, we do not expect most users to interact with these files. Instead, much of the data in the reference files is consolidated into specific extensions of the two primary DAP output files, the MAPS and model LOGCUBE files, as indicated in Figure 3. A complete description of the two main DAP output files is provided in Section 12 with the data models given in Appendix C.

Two of the six DAP modules, StellarContinuumModel and EmissionLineModel, employ a full-spectrum-fitting approach, and both of these modules use pPXF (Cappellari 2017); see Sections 7 and 9. The spectral templates used are critical to the pPXF procedure. The DAP repository

[39] M. Blanton; www.nsatlas.org.

[40] https://www.sdss.org/dr15/manga/manga-target-selection/ancillary-targets/

[41] When needed for covariance calculations (Section 6), the LOGRSS files are also used following a computation identical to what the DRP uses to produce the *griz* covariance matrices provided in DR15 (Aguado et al. 2019).

[42] https://www.sdss.org/dr15/manga/manga-data/data-model

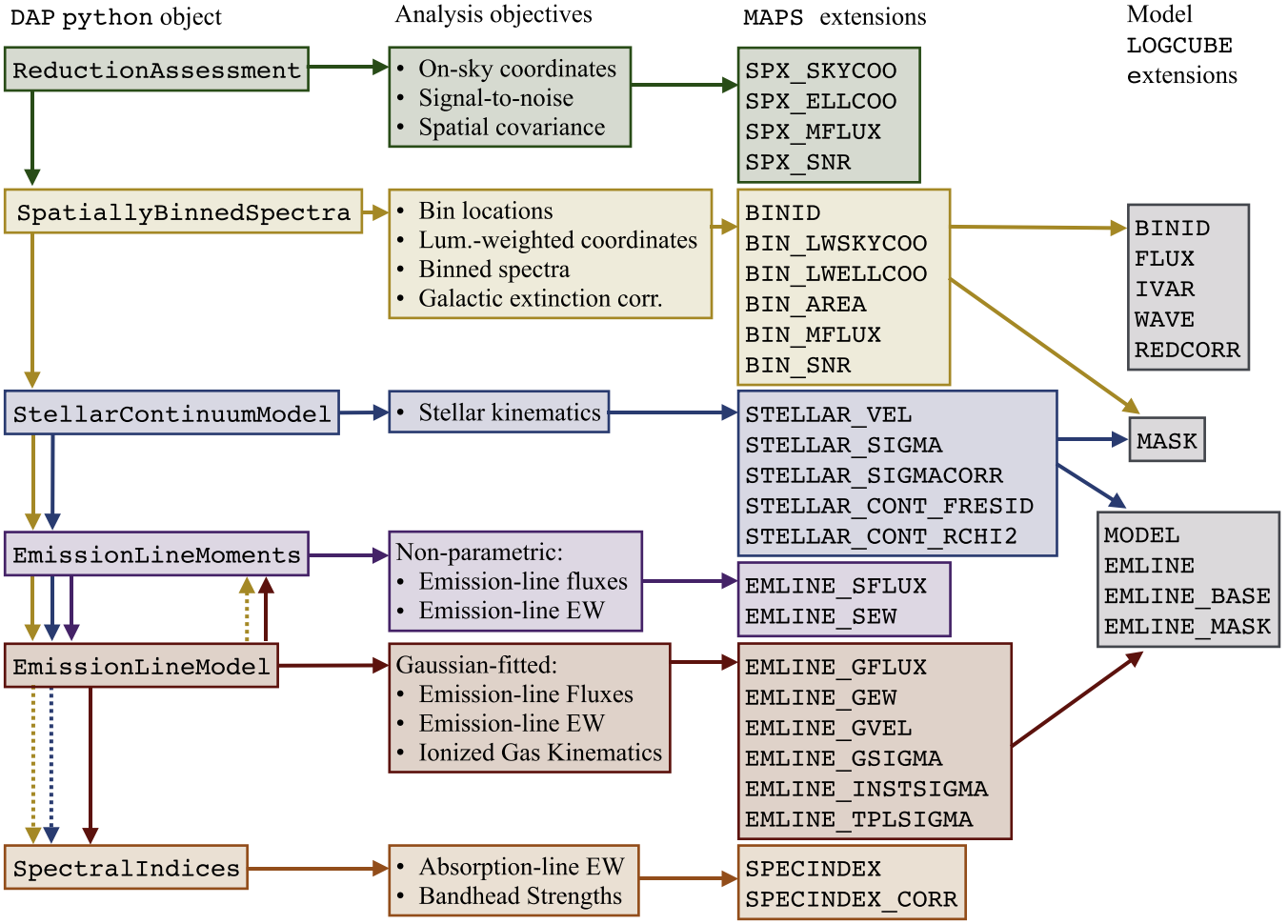

Figure 3. Schematic diagram of the DAP workflow. From left to right, the schematic provides the relevant Python modules, the analysis objectives of each module, and the associated MAPS and model LOGCUBE extensions generated by each module, as indicated by the arrows and colors. The Python modules, contained within the named DAP Python objects, are ordered from top to bottom by their execution order; an exception to this is that the emission-line moments are computed both before and after the emission-line modeling (see Sections 9 and 9.7), as indicated by the two sets of arrows pointing toward the EmissionLineMoments module. Arrow directions indicate the execution order and colors indicate the module dependencies. For example, the EmissionLineMoments object depends on the results of both the SpatiallyBinnedSpectra object and the StellarContinuumModel object. The dashed arrows indicate *conditional* dependencies. For example, the EmissionLineModel deconstructs the bins in the hybrid-binning approach, such that the SpectralIndices are independent of the primary results of the SpatiallyBinnedSpectra. However, there is an explicit dependence of the SpectralIndices on the SpatiallyBinnedSpectra when the hybrid-binning approach is not used.

provides a number of spectral-template libraries that we have collected over the course of DAP development (see Belfiore et al. 2019, Section 4.1); however, the results provided for DR15 focus on a distillation of the MILES (Sánchez-Blázquez et al. 2006; Falcón-Barroso et al. 2011) stellar-template library using a hierarchical clustering (HC) technique. We refer to the template library resulting from this analysis as the MILES-HC library, and we discuss the generation of this library in full in Section 5. Sections 6–10 discuss the details of the algorithms used in the six main DAP modules.

5. Hierarchical Clustering of a Spectral-template Library

To reduce computation time when using large stellar libraries as templates, one generally tries to select subsamples of stars that are representative of the entire library. For example, the execution time for the pPXF method, which the DAP uses to both measure the stellar kinematics (Section 7) and model the emission lines (Section 9), is typically slightly larger than $O(N_{\rm tpl})$ for $N_{\rm tpl}$ templates. Distillation of the information content of a spectral library into a minimal number of templates can therefore be critical to meeting the computational needs of large-scale surveys like MaNGA.

One way to sub-sample a library is to select stars that uniformly sample a grid in known stellar physical parameters, such as effective temperature ($T_{\rm eff}$), metallicity ([Fe/H]), and surface gravity (g) (e.g., Shetty & Cappellari 2015). The disadvantage of this approach is that stellar parameters may not always be available and are not necessarily a direct proxy of all relevant variation in the spectral information provided by the library. Alternatively, one could use a principle-component analysis to isolate the eigenvectors of the full stellar library (see Chen et al. 2012). However, one then loses the ability to enforce positivity constraints on the weights of the templates when modeling the galaxy spectra, which is a useful prior for reducing unphysical results at low S/N. Although more complex methods exist (e.g., Non-Negative Matrix Factorization, NMF; Lee & Seung 1999; see Blanton & Roweis 2007), we have adopted a simple approach that sufficiently avoids the

limitations of these alternatives and is generally applicable to any spectral-template library.

The key idea of our method is to apply a clustering algorithm (Jain et al. 1999) to the $N_{\rm tpl}$ spectral templates composed of M spectral channels by treating them as N vectors in an M-dimensional space.[43] For DR15, we have adopted a hierarchical-clustering approach (Johnson 1967), due to its simplicity, availability of robust public software, and the limited number of tuning parameters.[44] We have applied our approach to the full MILES stellar library[45] (Sánchez-Blázquez et al. 2006; Falcón-Barroso et al. 2011) of 985 stars by defining the distance between two spectra, S_j and S_k, as

$$d_{jk} = \frac{2\delta(S_j - S_k)}{\langle S_j \rangle}, \quad (1)$$

where $2\delta(S_j - S_k)$ is a robust estimate of the standard deviation of the residual, computed as one half of the interval enclosing 95.45% of the residuals, in a `pPXF` fit of spectrum S_k using spectrum S_j as the template. For this exercise, we include an eighth-order additive Legendre polynomial in the `pPXF` fits to be consistent with the method used when fitting the stellar kinematics of the galaxy spectra in the `DAP` (Section 7). The individual elements of d_{jk} from Equation (1) are used to construct a distance matrix for input to a hierarchical-clustering algorithm.[46]

We form flat clusters such that the cluster constituents have a maximum distance of $d_{\rm max}$; lower values of $d_{\rm max}$ yield a larger number of flat clusters. To construct the spectral templates for the distilled library, we normalize each MILES spectrum to a mean of unity and we then average all of the spectra in each cluster without weighting. Throughout this paper, we refer to the result of our hierarchical clustering of the MILES stellar library as the `MILES-HC` library. We have optimized $d_{\rm max}$ by comparing `pPXF` fits of high-S/N MaNGA spectra from a few representative young/old galaxies using either the full set of 985 MILES stars or the `MILES-HC` library produced by the given iteration of $d_{\rm max}$.

For $d_{\rm max} = 0.05$,[47] our analysis yields 49 clustered spectra from the full MILES stellar library of 985 spectra. The number of spectra assigned to each cluster varies dramatically, from clusters composed of individual spectra to others that collect hundreds of stars from the MILES library. However, as one would expect, the clusters tend to concentrate in regions of stellar parameter space, as shown in Figure 4. Although compelling from a perspective of stellar spectroscopy, the details of the distribution are less important to our application than whether or not the clustering has successfully captured the information content of the full MILES spectral library relevant to our full-spectrum fitting.

[43] One can think of a vast range of practical implementations of this general idea, given the large number of solutions that were proposed for the clustering problem; however, our simple approach has proven reasonable for our purposes, if not necessarily optimal.

[44] The Python code that we have used to generate the library is available at https://github.com/micappe/speclus.

[45] We used MILES V9.1, which is available at http://miles.iac.es/.

[46] Specifically, we use the `scipy` (Jones et al. 2001) function `cluster.hierarchy.linkage` with `method='average'`. This implements the nearest-neighbors chain hierarchical-clustering algorithm described by Müllner (2011).

[47] Specifically, we use the `scipy` function `cluster.hierarchy.fcluster` with `criterion='distance'` and a `threshold` of 5%.

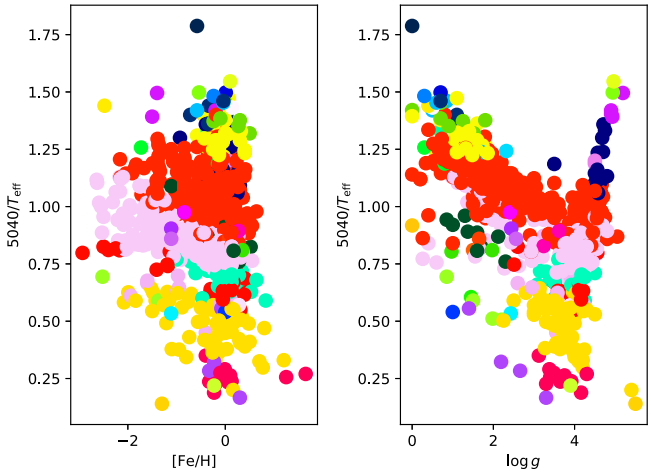

Figure 4. Effective temperature $T_{\rm eff}$, metallicity [Fe/H], and surface gravity g of the stars in the MILES stellar library (Falcón-Barroso et al. 2011). Each datum is assigned a color based on its assigned cluster from the hierarchical-clustering algorithm (Section 5). Some clusters contain a single star, whereas others include about a hundred stars. Cluster boundaries generally do not follow lines of constant stellar parameter due to the degeneracy between the three physical parameters.

From the original set of 49 cluster spectra, we remove templates with prominent emission lines or relatively low S/N (e.g., from clusters composed of a single spectrum), leading to a final set of 42 spectra in the `MILES-HC` template library, as shown in Figure 5.[48] We compare the stellar kinematics measured using the full MILES and `MILES-HC` libraries in Section 7.3. As one would expect, the use of `MILES-HC` yields a moderately worse fit, as determined by the fit residuals and chi-square statistics; however, the affect on the resulting kinematics is acceptable, particularly given the gain of roughly a factor of 25 in execution time. In their Section 4, Belfiore et al. (2019) also compare the emission-line modeling results when the stellar continuum is fit using the `MILES-HC` library and various simple-stellar-population (SSP) templates. The `MILES-HC` library shows specific differences in the continuum shape and Balmer absorption depths compared to the Bruzual & Charlot (2003, BC03) library, given its lack of early-type (O) stars. However, the quality of the fits to the MaNGA spectra using `MILES-HC` are generally no worse than when using SSP templates.

6. Spatial Binning

Unbiased measurements of stellar kinematics require a minimum S/N, particularly for the stellar velocity dispersion. It is therefore generally necessary to bin spectra by averaging neighboring spaxels to meet a given S/N threshold. To bin for this purpose, we use the adaptive spatial-binning scheme implemented by the Voronoi algorithm of Cappellari & Copin (2003).[49] The data cube construction scheme in MaNGA (Law et al. 2016, Section 9) follows the method of Shepard (1968) (see also Sánchez et al. 2012), leading to significant covariance between adjacent spaxels that must be accounted for when combining spaxel data. Indeed, given that the Voronoi-binning algorithm is predicated on meeting a minimum S/N, the

[48] These spectra are made available through the `DAP` GitHub repository; specifically, https://github.com/sdss/mangadap/tree/master/data/spectral_templates/miles_cluster.

[49] In DR15 specifically, we use the Python package `vorbin` version `3.1.3` found at https://pypi.org/project/vorbin/.

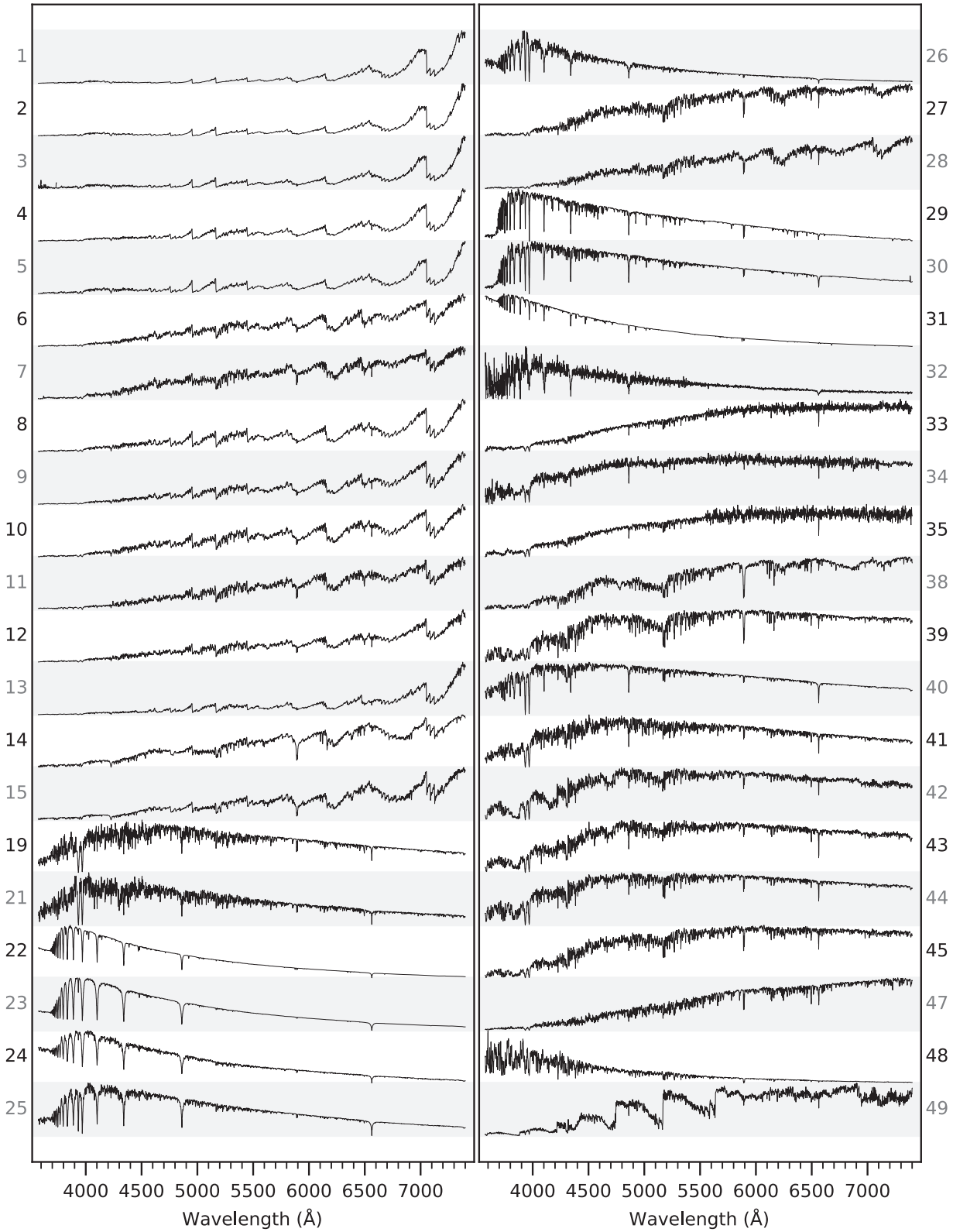

Figure 5. The 42 templates in the MILES-HC library identified along the left or right ordinate by their cluster group number. Missing clusters in the sequence (e.g., cluster 16) were removed because of low S/N or the presence of emission lines from flaring late-type stars. The spectra are ordered by their cluster number, which is not generally related to the mean stellar parameters of the cluster constituents.

success of the algorithm hinges on an accurate calculation of the binned S/N. However, calculation of the full covariance matrix in each data cube is prohibitively expensive, prompting a few approximations in our approach.

The following sections describe the first two modules of the DAP workflow (Figure 3) that ultimately yield the binned spectra used for the determination of the stellar kinematics. The distinction between these two modules is that the first is

independent of any specific binning algorithm (Sections 6.1 and 6.2), whereas the second performs the binning itself (Sections 6.3 and 6.4). The incorporation of spatial covariance when aggregating spaxels to meet a minimum S/N (Section 6.3) and when propagating the uncertainties in the binned spectra (Section 6.4) are treated separately for computational expediency.

6.1. On-sky Spaxel Coordinates and Data Cube Mask

The DRP provides a World Coordinate System (WCS) for each data cube, which the DAP uses to calculate the on-sky coordinates of each spaxel relative to the target center. The target center is provided in the data cube header with the keywords OBJRA and OBJDEC.[50] The on-sky coordinates provided by the DAP are sky-right in arcseconds, with positive R.A. offsets toward the East; note that the abscissae in Figure 1 increase from right to left. The DAP then uses the photometric position angle and ellipticity to calculate the semimajor-axis coordinates, R and θ. For DR15, these are simply calculated and included in the output MAPS file.

The DRP also provides detailed masks for each wavelength channel, which the DAP uses to exclude measurements from analysis in any given module. Large swaths of the full MaNGA spectral range can be masked by the DRP because of broken fibers, known foreground-star contamination, detector artifacts, or (in the majority of cases) simply because the spaxel lies outside of the hexagonal IFU FOV. The DAP excludes measurements affected by these issues by ignoring any measurement flagged as either DONOTUSE or FORESTAR. For each spaxel, we calculate the fraction of the MaNGA spectral range, $\delta\Lambda$, that is viable for analysis. For DR15, the DAP ignores any spaxel with $\delta\Lambda < 0.8$. As a representative example, Figure 6 shows the viable fraction of spaxels, $\delta\Omega$, with *any* valid flux measurement as a function of $\delta\Lambda$ for the data cubes observed by plate 7495.

6.2. Spectral S/N and Spatial Covariance

Both as a basic output product and for binning purposes, the DAP calculates a single measurement of S/N for each spaxel. In DR15, this fiducial S/N—hereafter referred to as S/N_g–is the average S/N per wavelength channel weighted by the g-band response function.[51] We calculate S/N_g for all spaxels, excluding masked channels, regardless of whether or not they meet our criterion of $\delta\Lambda > 0.8$ (Section 6.1).

For this fiducial S/N_g, we also calculate a single spatial covariance matrix in two steps. (1) We calculate the spatial correlation matrix for the wavelength channel at the response-weighted center of the g-band following Equation (7) from Law et al. (2016, see Equation (2) herein). We find that the spatial correlation matrix varies weakly with wavelength over the g-band such that, to first order, we can simply adopt the correlation matrix from this single wavelength channel. (2) We renormalize the single-channel correlation matrix by the mean variance in the flux over the g-band to construct a covariance matrix.

[50] These are typically, but not always, the same as the pointing center of the IFU given by the keywords IFURA and IFUDEC.

[51] Specifically, we use the response function produced by Jim Gunn in 2001, which is provided at https://www.sdss.org/wp-content/uploads/2017/04/filter_curves.fits with the description the SDSS Survey imaging camera at https://www.sdss.org/instruments/camera/.

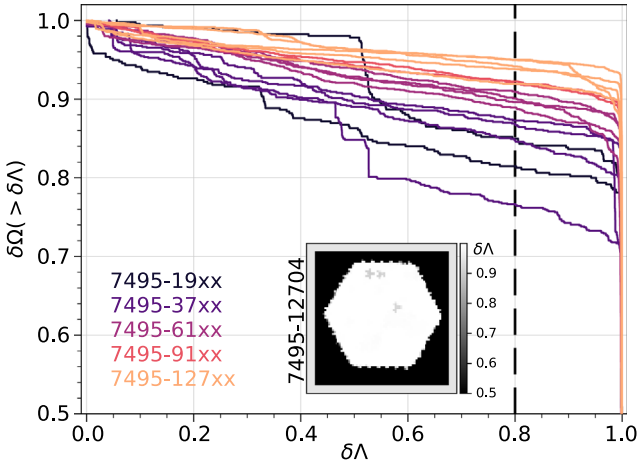

Figure 6. The fraction of valid wavelength channels ($\delta\Lambda$) over the full spectral range vs. the fraction of spaxels ($\delta\Omega$; see the definition in Section 6.1) with at least $\delta\Lambda$ over the IFU field of view. Data are shown for all 17 observations from plate 7495, colored by the IFU size. The inset map shows $\delta\Lambda$ in each spaxel of the data cube for observation 7495–12704: the hexagonal area with non-zero $\delta\Lambda$ is surrounded by a buffer of spaxels with $\delta\Lambda = 0$ resulting from the data cube construction. The DAP only analyzes spaxels with $\delta\Lambda > 0.8$ (vertical dashed line).

Figure 7 provides the correlation matrix for wavelength channel 1132 in data cube 8249-12705 calculated following the first step described above, where the correlation coefficient is defined as $\rho_{jk} = C_{jk}/\sqrt{C_{jj}C_{kk}}$ and C_{jk} is the covariance between spaxels j and k. Only spaxels with $S/N_g > 1$ are included in the figure. Critically, note that the indices j and k are *not* the two-dimensional indices of an individual spaxel on sky but are instead indices for the spaxels themselves. In other words, spaxel j will have on-sky coordinates (x_j, y_j) and appropriate array indices in the DAP map. This explains the diagonal banding in Figure 7 as an effect of spatially adjacent spaxels being separated by the width of the map in one-dimension in the correlation matrix. Figure 7(b) is an expanded view of the ± 20 pixels about the main diagonal and has a width of approximately 10 pixels. The number of discrete diagonal bands in Figure 7(a) and the width of the off-diagonal distribution in Figure 7(b) demonstrates that spaxels separated by fewer than 5 or 6 spaxels have $\rho > 0$, which is consistent with the subsampling of the MaNGA 2.″5-diameter fiber beam into $0.''5 \times 0.''5$ spaxels.

We show this explicitly in Figure 8, which combines the correlation data for all spaxels within the convex hull of the fiber centers used to construct wavelength channel 1132.[52] We find in this case, and generally, that ρ_{jk} is well-fitted by a Gaussian distribution in the distance between spaxels, D_{jk}. The optimal fit to this channel is given in the figure, where the Gaussian has a scale parameter of $\sigma = 1.92$ spaxels ($0.''96$).

Finally, as a metric for the degree of covariance in this wavelength channel, we compute $N^2/\sum_{jk}\rho_{jk} = 142.5$, which provides a rough estimate of the number of independent measurements. Note that in the limit of fully independent and fully correlated measurements, $N < \sum_{jk}\rho_{jk} < N^2$, respectively. As expected, the rough estimate of independent measurements within the data cube is comparable to the

[52] Spaxels outside the convex hull of the fiber centers naturally have larger correlation coefficients because fewer fibers contribute to these spaxels. More generally, this effect will be true of regions in the IFU FOV masked at the fiber level, due to, e.g., a broken or very low-throughput fiber.

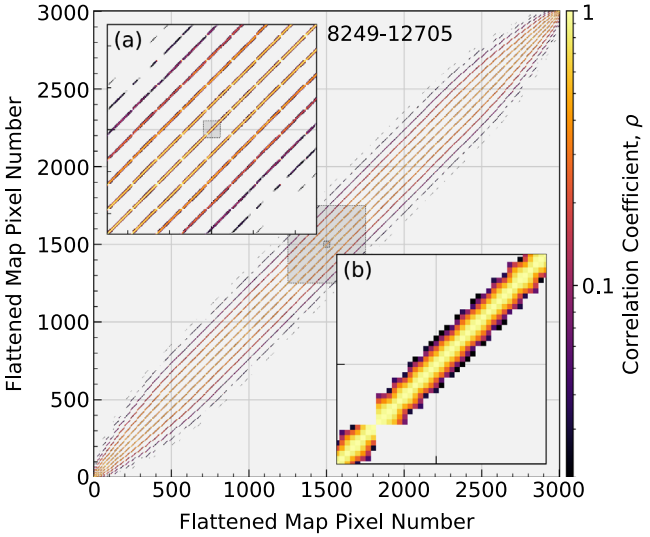

Figure 7. The correlation matrix in channel 1132 ($\lambda = 4700$ Å) of data cube `8249-12705` for all spaxels with $S/N_g > 1$. The correlation matrix is symmetric and has a correlation coefficient of $\rho = 1$ along the matrix diagonal, by definition. The majority of the matrix is empty with $\rho = 0$. The inset panels provide an expanded view of two matrix subregions: panel (a) shows ± 250 pixels around the matrix center as indicated by the large gray box, and panel (b) shows ± 20 pixels around the matrix center as indicated by the small gray box, also shown in panel (a). The diagonal banding in panel (a) is an artifact of the ordering of adjacent pixels in the flattened vector of the data cube spatial coordinates; adjacent pixels are separated by the width of the map in one on-sky dimension. The number of bands in panel (a) roughly matches the width in pixels along the main diagonal with non-zero ρ in panel (b), as expected by the spatial correlation acting along both on-sky dimensions.

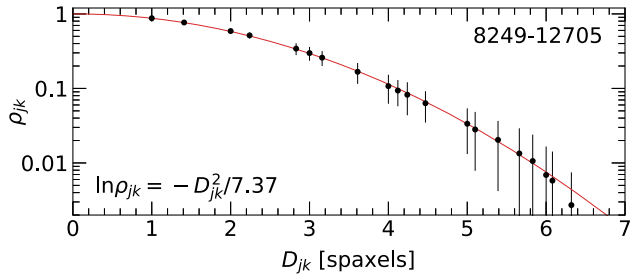

Figure 8. The mean (points) and standard deviation (errorbars) of the correlation coefficient, ρ_{jk}, for all spaxels within the convex hull of the fiber-observation centers at channel 1132 ($\lambda = 4700$ Å) of data cube `8249-12705` as a function of the spaxel separation, D_{jk}. The best-fitting Gaussian trend (red) has a scale parameter of $\sigma = 1.92$ spaxels, leading to the equation provided in the bottom-left corner.

number of fibers in the relevant IFU (127); however, this is substantially smaller than the 1905 independent fiber observations used to construct the data cube. For a more in-depth discussion of data cube reconstruction and a method that minimizes data cube covariance, see Liu et al. (2019).

6.3. Voronoi Binning with Covariance

As we have stated above, the fidelity of the Voronoi-binning approach critically depends on a proper treatment of the spatial covariance. For illustration purposes, we have applied the Voronoi-binning algorithm to the S/N_g measurements for data cube `8249-12705` both with and without an accounting of the spatial covariance. A map and radial profile of the S/N_g measurements are shown in the top-left and bottom panels of Figure 9, respectively, and the correlation matrix used is shown in Figure 7. Although the threshold used in DR15 is $S/N_g \sim 10$, we use a threshold of $S/N_g \sim 30$ to accentuate the effect. Application of the algorithm without using the correlation matrix data results in the bin distribution shown in the upper-middle plot of Figure 9; the distribution resulting from the formally correct S/N_g calculation is shown in the upper-right panel.

The effect of the covariance dramatically increases the number of spaxels needed to reach the target S/N_g, as evidenced by comparing the size of the bins in the upper-middle and upper-right panels. If we apply the formal calculation of the S/N_g to the bins generated without the covariance (red points in the bottom panel of Figure 9), we find that the actual S/N_g of these bins is far below the desired threshold. In this example, we note that the calculation that includes covariance also falls short of the target S/N_g; however, this is due to the details of Voronoi-binning algorithm and is not an inconsistency in the S/N calculation. In detail, not every S/N function can partition the FOV into compact bins with equal S/N. To increase their S/N in this example, the outermost bins would have to become elongated (e.g., like a circular annulus following the edge). This is prevented by the roundness criterion of the Voronoi-binning algorithm and, therefore, limits the S/N_g of these bins. However, this example is not representative of a systematic difference between our target $S/N_g \sim 10$ and what is achieved by our use of the Voronoi-binning algorithm (see Figure 26).

To minimize the systematic errors at low S/N_g for the stellar velocity dispersions, we have chosen a S/N_g threshold of 10 per wavelength channel for DR15, which is discussed further in Section 7. This is sufficient for the first two kinematic moments, but one likely needs an increased threshold for higher order moments (h_3 and h_4).

6.4. Spectral Stacking Calculations

For use in the subsequent modules of the `DAP`, the procedure used to stack spaxels must yield the flux density, inverse variance, mask, and wavelength-dependent spectral resolution of each binned spectrum. The stacked flux density is a simple masked average of the spectra in each bin, whereas the computations for the uncertainty and spectral resolution of the binned spectra are more subtle and are discussed in detail below. These procedures are fundamentally independent of the specific algorithm that determines which spaxels to include in any given bin, and our treatment of spatial covariance is slightly different.

The variance in the binned spectra is determined from the covariance matrix as follows. Similar to the calculation of the covariance in the data cubes, the covariance in the binned spectra at wavelength λ is

$$\boldsymbol{C}_{\lambda,\mathrm{bin}} = \boldsymbol{T}_{\mathrm{bin}} \boldsymbol{C}_{\lambda,\mathrm{spaxel}} \boldsymbol{T}_{\mathrm{bin}}^{\top}, \quad (2)$$

where $\boldsymbol{C}_{\lambda,\mathrm{spaxel}}$ is the covariance matrix for the spaxel data and $\boldsymbol{T}_{\mathrm{bin}}$ is an $N_{\mathrm{bin}} \times N_{\mathrm{spaxel}}$ matrix where each row flags the spaxels that are collected into each bin. To avoid the expensive calculation of the full data cube covariance matrix, Law et al. (2016)—following the original proposal by Husemann et al. (2013)—recommended the easier propagation of the error that ignores covariance and provided a simple functional form for a factor, f_{covar}, that nominally recalibrates these error vectors for

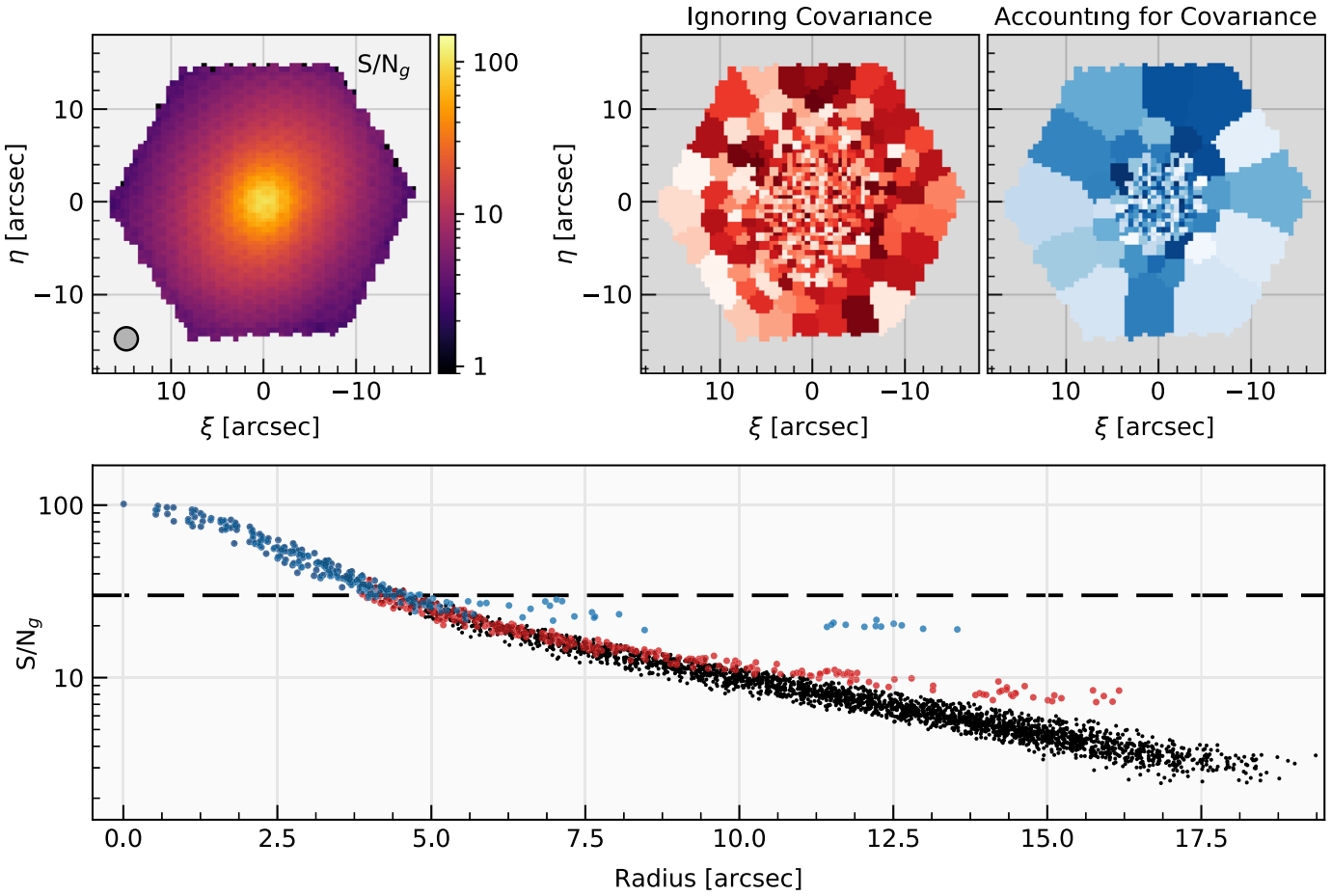

Figure 9. Effect of spatial covariance on the result of the Voronoi binning algorithm. The top-left panel shows the S/N$_g$ measurements for data cube 8249-12705. We then apply the Voronoi binning algorithm to these data with a S/N$_g$ threshold of 30. The resulting bin distribution that does not include the spatial correlation from Figure 7 is shown in the top-middle panel, and the bin distribution that does include the correlation is shown in the top-right panel. The colors in the top-middle and top-right panels are used to differentiate between spaxels in a given bin. The bottom panel shows the formally correct S/N$_g$ as a function of radius for the individual spaxels (black), the bins derived assuming no covariance (red), and the bins that include the covariance (blue). The Voronoi algorithm expects the red data to have S/N$_g$ ∼ 30 based on the S/N$_g$ calculation that excludes covariances; however, the formally correct S/N$_g$ is well below that.

the effects of covariance based on the number of binned spaxels. In the DAP, we instead base $f_{\rm covar}$ on directly calculated covariance matrices sampled from 11 wavelength channels across the full spectral range of the data. As an example, Figure 10 shows the applied recalibration for the Voronoi bins in observation 8249–12705 compared to our suggested nominal calibration provided by Law et al. (2016).

It is important to note that covariance persists in the rebinned spectra, even between large, adjacent bins.[53] Therefore, it is important to account for this covariance if users wish to rebin the binned spectra;[54] however, in this case, we recommend simply rebinning the original data cube.

The spectral resolution in the binned spectrum is determined by a nominal propagation of the per-spaxel measurements of the line-spread function (LSF), newly provided with the data cubes released in DR15 (Aguado et al. 2019). Similar to how these LSF cubes are produced by the DRP, we calculate the second moment of the distribution defined by the sum of the

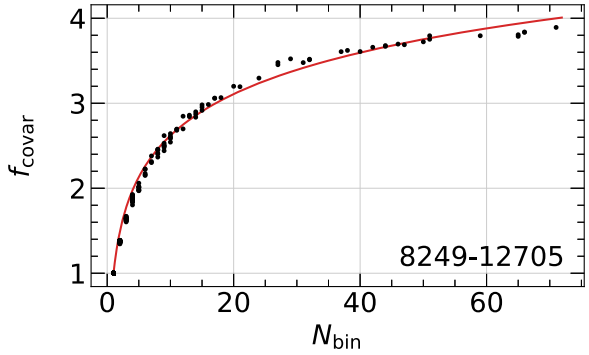

Figure 10. The computed factor that properly rescales noise vectors computed without accounting for covariance to those that do, $f_{\rm covar}$, as a function of the number of binned spaxels, $N_{\rm bin}$, for the Voronoi bins constructed for observation 8249–12705 (points). These data are based on the median ratio obtained from the direct calculation of the covariance matrix in 11 wavelength channels spanning the full spectral range of the data. For comparison, the nominal calibration, $f_{\rm covar} = 1 + 1.62\log(N_{\rm bin})$, from Law et al. (2016) is shown in red.

Gaussian LSFs determined for each spaxel in the bin; i.e.,

$$\sigma^2_{\rm inst,bin}(\lambda) = \frac{1}{N_{\rm bin}} \sum_{i}^{N_{\rm bin}-1} \sigma^2_{{\rm inst},i}(\lambda), \quad (3)$$

[53] It is effectively impossible to rebin the MaNGA *data cube* in a way that removes the covariance. One has to restart with the fiber data in the RSS files.
[54] If needed, the 11 covariance matrices used to recalibrate the error in the binned spectra are provided in the DAP reference file.

where $\sigma^{-1}_{\mathrm{inst},i} = R_i\sqrt{8\ln 2}/\lambda$ for each spaxel i, with resolution R_i, in the bin. A limitation of this calculation is that the assumption of a Gaussian LSF with a dispersion of $\sigma^2_{\mathrm{inst,bin}}$ for the binned spectrum becomes less accurate as the range in $\sigma_{\mathrm{inst},i}$ increases. For MaNGA data, the variation in the LSF between spaxels in a data cube is generally only a few percent, meaning that this should not be a concern to first order. However, this may be more of an issue when binning spectra across observations. Additionally, deviations of the LSF from a Gaussian *will* be a concern for, e.g., higher-order moments of the stellar line of sight velocity distribution (LOSVD) (e.g., h_3, h_4). In DR15, we only provide the first two moments of the stellar LOSVD, but this could be a concern for future releases and certainly for those performing their own measurements of h_3, etc.

Although never used in the DAP directly, we also calculate the luminosity-weighted coordinates of the binned data in the MAPS output file. These are simple weighted means of the coordinates of each spaxel using the mean g-band flux (also provided in the MAPS file) as the weight.

6.5. Galactic Extinction Correction

After the stacking procedure, all spectra to be fitted are corrected for Galactic extinction using the $E(B-V)$ value from the Schlegel et al. (1998) maps provided by the EBVGAL keyword in the header of the DRP LOGCUBE file. In DR15, we use the Galactic extinction law presented by O'Donnell (1994); the exact de-reddening vector used for each data cube is provided in the model LOGCUBE output file (Section 12.2).

6.6. Flagging

The S/N_g metrics and binned spectra are flagged according to the following criteria. Spaxels that are ignored because they do not meet the S/N_g or spectral-coverage criteria are masked as IGNORED. Any binned spectrum with individual channels that were masked for *all* spectra in the bin are masked as FLUXINVALID, and pixels with invalid inverse variance values are flagged as IVARINVALID; see Table 9.

7. Stellar Kinematics

The workhorse of our stellar-continuum module (discussed here) and emission-line fitting module (Section 9; Belfiore et al. 2019) is the penalized pixel-fitting method[55] (pPXF) by Cappellari (2017), which is an upgrade to the original algorithm by Cappellari & Emsellem (2004). We refer the reader to these papers for detailed descriptions of the method. In brief, pPXF assumes that a galaxy spectrum is composed of a mixture of template spectra, convolved with the LOSVD function of the kinematic component to which each is assigned. The primary improvement introduced by Cappellari (2017) is that this convolution is now accurate to an arbitrarily small velocity dispersion, which is achieved by defining the convolution kernel in Fourier space (see Section 9.1.1; Equations (14) and (15)). Currently, the DAP assigns all stellar-continuum templates to a single kinematic component, while the emission-line fitting module allows for multiple dynamical components (see Johnston et al. 2013; Mitzkus et al. 2017). The DAP also includes low-order additive and/or multiplicative adjustments to the continuum via Legendre polynomials, as allowed by pPXF functionality.

The core pPXF algorithm is abstracted and generalized to allow for its broad application; therefore, it is important to discuss its specific use in the DAP for analyzing MaNGA spectra. Other applications of the pPXF method and software for kinematic measurements in IFS galaxy surveys include SAURON (Emsellem et al. 2004), ATLAS^{3D} (Cappellari et al. 2011), VENGA (Blanc et al. 2013), CALIFA (Falcón-Barroso et al. 2017), and SAMI (van de Sande et al. 2017; Scott et al. 2018).

Our primary concerns when optimizing our approach are: (1) the selection of the template spectra (see Sections 5, 7.3, and 7.4.1), both in terms of their pedigree (empirical versus theoretical and individual stars versus stellar-population synthesis) and their resolution and sampling; (2) the limitations in the results caused by the instrumental LSF and S/N; and (3) the optimization of the parameters provided by the algorithm, such as the order of the polynomials included in the fit and the penalization bias applied during fits that include the non-Gaussian moments (h_3, h_4) of the LOSVD.

In this section, we describe our fitting algorithm as it has been applied to the data cubes provided in DR15 (Section 7.1). We briefly compare the stellar kinematics measured using the full MILES and MILES-HC libraries to motivate our use of the latter (Section 7.3) and we justify some of the nuances of our fitting algorithm (Section 7.4) based on the data briefly discussed in Section 7.2. We quantify the performance of our algorithm using both simulated and observed MaNGA data (Section 7.5), and we describe the quality flags provided in DR15 (Section 7.6). Finally, given our particular treatment and presentation of the stellar velocity-dispersion measurements, we provide guidance and recommendations for their use in Section 7.7.

7.1. Fitting Algorithm

7.1.1. Template Preparation

To be flexible to changes in the data and to facilitate testing, we have built methods into the DAP that allow us to easily switch between spectral template libraries in our full-spectrum-fitting modules. The template libraries provided with the DAP repository[56] may be linearly or logarithmically sampled in wavelength, can be provided with a vacuum or air wavelength calibration, have different spectral resolutions, and adopt different conventions for their overall flux normalization. For example, the MILES-HC spectra have a spectral resolution of $\Delta\lambda = 2.5$ Å (Beifiori et al. 2011; Falcón-Barroso et al. 2011) and a normalization near unity, whereas the M11-STELIB spectra, a set of stellar population models provided by Maraston & Strömbäck (2011), have a spectral resolution of $\Delta\lambda = 3.4$ Å and a normalization in physical units of erg s^{-1} Å^{-1} $\mathcal{M}_\odot^{-1}$. Therefore, we have designed the DAP to be flexible to this variety among the spectral libraries by always performing a few steps to prepare the templates for use when fitting the MaNGA spectra. Here, we focus on the template set used for DR15, the MILES-HC library (Section 5); however, in Belfiore et al. (2019), we use many more template libraries to test the effects of the stellar-continuum modeling on the emission-line properties.

[55] We use the Python package ppxf version 6.7.8 found here: https://pypi.org/project/ppxf/.

[56] https://github.com/sdss/mangadap/tree/master/data/spectral_templates

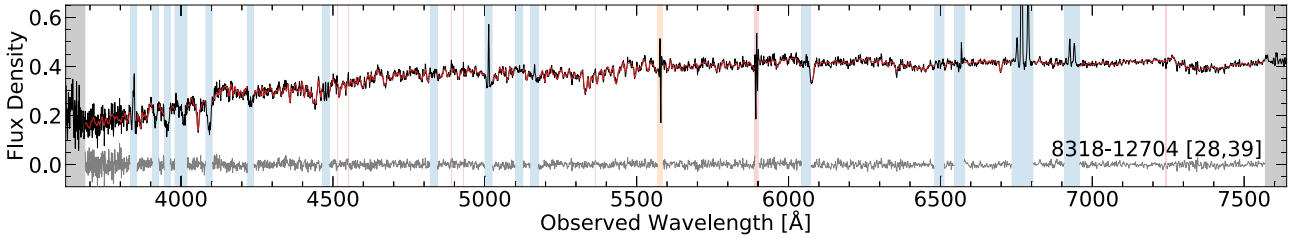

Figure 11. An illustration of the spectral masking both input to and resulting from a typical execution of the stellar-continuum modeling used to measure stellar kinematics. The observed spectrum (black) is from observation `8318-12704`, specifically the spaxel at coordinates $(x, y) = (28, 39)$. (If read using `astropy.io.fits`, this is the spectrum in the `FLUX` array selected using the slice `[:,39,28]`.) Before beginning the fit, masks are constructed to omit regions outside the spectral range of the template spectra (vertical gray bands), at the expected locations of emission lines (vertical blue bands), and at the location of the strong [O I] night-sky emission line (vertical orange band). After the first iteration, 3σ outliers are identified and rejected, which in this case omits a few individual channels and an artifact at $\lambda \sim 5900$ Å (vertical red bands). The resulting model is shown in red, overlayed on the observed spectrum, and the model residuals are shown in gray.

The template preparation steps are as follows: (a) the wavelengths are converted to vacuum, if necessary, to match the MaNGA data; (b; *optional*) we nominally match the template-library resolution to the MaNGA data by convolving each spectrum with a wavelength-dependent Gaussian kernel (Appendix A), (c) we resample each template to a spectral channel width that is a fixed fraction of the MaNGA `LOGCUBE` spectral sampling;[57] and (d) we normalize the flux such that the mean flux *over all templates* is unity. In our nominal approach to measuring stellar kinematics, we skip step (b), leaving the spectral resolution of the templates unaltered. This offset in spectral resolution between the template and galaxy spectra is particularly important: the implications for our fitting algorithm are discussed in Section 7.1.5 and the reasoning behind this choice is discussed in Section 7.4.3. The spectral sampling of the logarithmically binned MaNGA spectra is fixed to $\delta \log \lambda = 10^{-4}$, which corresponds to a velocity scale of $\Delta V \approx 69$ km s^{-1}. We take advantage of the optional behavior of `pPXF` in allowing the template spectra to be sampled at some integer rate higher than the object spectra to avoid under-sampling high-resolution libraries. Therefore, our nominal approach in step (c) is to sample the `MILES-HC` library template spectra four times per MaNGA spectral channel ($\delta \log \lambda = 2.5 \times 10^{-5}$). However, we note that Cappellari (2017, Figure 2) demonstrates this oversampling of the template spectra is not strictly necessary to accurately recover velocity dispersions for an LOSVD that is under-sampled.

7.1.2. Masking

A detailed mask is constructed for each spectrum in the MaNGA data cube before passing the spectrum to `pPXF`, as illustrated in Figure 11: first, `pPXF` restricts the number of spectral channels in the template spectra (accounting for any pixel-scale differences) to be the same or larger than in the object spectra. However, MaNGA spectra typically have a larger spectral range than the templates we have included in our testing. In particular, the `MILES-HC` library covers ~3600–7400 Å, which is little more than half of the MaNGA spectral range (see Figure 2). The `DAP` therefore censors the MaNGA spectra based on the expected overlap with the Doppler-shifted template spectra, given z_0 and an assumed maximum velocity range of ± 400 km s^{-1}. We mask an additional $\pm 3\sigma_*$, where σ_* is the stellar velocity dispersion, at the edges of the spectral range to limit convolution aliasing. Instead of dynamically masking during the fitting, we mask assuming a maximum of $\sigma_* = 400$ km s^{-1}. Both of these masks are shown in gray in Figure 11. Second, we mask any spectrum with $S/N_g < 1$, any spectral channels with `MANGA_DRP3PIXMASK` bits set to either `DONOTUSE` or `FORESTAR` by the `DRP`, and the region from 5570–5586 Å to avoid the near-ubiquitous subtraction residuals of the strong [O I] night-sky line; the latter is shown in orange in Figure 11. Third, we mask a ± 750 km s^{-1} region around the expected center of each emission line in Table 3, Doppler shifted to match the expected recession velocity using z_0 (blue in Figure 11). The emission-line mask is applied regardless of whether or not any emission line is detected. This emission-line mask is generally sufficient for MaNGA galaxies; however, particularly broad emission-line cores (e.g., broad-line active galactic nucleus (AGN) spectra) or wings (e.g., star formation outflows) are notable exceptions (see our discussion of performance in Section 11).

7.1.3. Fit Iterations

After preparing the templates and constructing the default mask, the `DAP` proceeds through two fit iterations. Each fit iteration uses a common set of templates, an additive eighth-order Legendre polynomial, and a Gaussian LOSVD; the necessary vetting and optimization required for fitting a Gauss–Hermite LOSVD has not yet been done for MaNGA within the `DAP`. Each iteration runs `pPXF` twice, once to isolate 3σ outliers in the spectrum and then with the outlying wavelength channels removed. The 3σ outliers are determined using a 100-channel (~6900 km s^{-1}) boxcar determination of the local mean and standard deviation in the fit residuals.

The first iteration fits the masked average of all spectra in the data cube (i.e., there is only one spectrum fit during this iteration) to isolate the subset of templates allocated non-zero weight. All spectra, either from a spatial bin or individual spaxel, are fit in the second iteration. Templates not included in the non-negative least-squares algorithm used by `pPXF` to solve for the template weights in the first iteration are excluded from the second iteration. This downsampling of the templates used in the second iteration both expedites that iteration and limits the effect of noise-driven inclusion of templates in fits to lower S/N_g spectra. We discuss the effects of limiting the templates used to fit each spectrum on the resulting kinematics in Section 7.4.1. The output `MAPS` files (Section 12; Appendix C) provide the measurements from the second iteration.

[57] This resampling allows for a non-uniform wavelength step as a function of wavelength, primarily to account for the nonlinear conversion from air to vacuum wavelengths in step (a).

7.1.4. Velocity Definition

The DAP does not de-redshift the spectra before executing the fits used to determine the stellar kinematics or when performing the emission-line modeling (Section 9). However, the velocities provided in the MAPS files have been offset to remove their cosmological redshift with respect to the solar barycentric rest frame. This is done in two steps. First, the velocities returned by pPXF are (see Section 2.3 of Cappellari 2017 for an explanation)

$$\begin{aligned} V_{\rm ppxf} &= dV\Delta p \\ &= c\ln(\lambda_{\rm obs}/\lambda_0) \\ &= c\ln(1+z) \end{aligned} \quad (4)$$

where $dV = c\ d\ln\lambda$ is the size of the logarithmically binned pixel in km s^{-1} and Δp is the shift in pixels found to attain the best fit between the template and galaxy spectra. Given that we are not deredshifting the spectra before executing pPXF, we must first use Equation (4) to calculate the observed redshift,[58] $z_{\rm obs}$, from $V_{\rm ppxf}$ for each spaxel. Second, for each galaxy, we remove from these observed redshift measurements the effect of the *input* cosmological redshift of the galaxy, z_0, to bring the stellar velocities to the reference frame of the galaxy. These are the values that satisfy standard Newtonian laws that one should use for, e.g., dynamical models (see Section 2.4 of Cappellari 2017 for an explanation), and these are the velocities (see Hogg 1999) reported in the output MAPS file:

$$V = c\ (z_{\rm obs} - z_0)/(1+z_0). \quad (5)$$

The cosmological redshift, z_0, is most often identical to the redshift provided by the NSA, except when NSA data is not available for the galaxy (e.g., it is an ancillary target) or when the NSA redshift has been corrected based on an improved measurement from the MaNGA data itself (e.g., Talbot et al. 2018).

7.1.5. Velocity-dispersion Corrections

As discussed above, the approach of the DAP is to fit the templates to the MaNGA spectra at their native resolution (see Section 7.4.3). This means that the velocity dispersion returned by pPXF will be

$$\sigma^2_{\rm obs} = \sigma^2_* + \delta\sigma^2_{\rm inst}, \quad (6)$$

where σ_* is the true astrophysical stellar velocity dispersion and $\delta\sigma_{\rm inst}$ is an effective difference in the instrumental dispersion of the template and MaNGA data. It is useful to note that even in the case where one obtains template spectra from the same instrument as the galaxy data, velocity-dispersion corrections may still be necessary given the redshift of the galaxy spectra; see the Appendix of Westfall et al. (2011) and additional discussion below.

For DR15, we provide a first-order estimate of $\delta\sigma_{\rm inst}$ based on the average quadrature difference in the instrumental dispersion of the template and object spectra over the region fit by pPXF. That is, we calculate

$$\delta\sigma^2_{\rm inst} = \frac{c^2}{8\ln(2)N_{\rm fit}} \sum_{i=0}^{N_{\rm fit}-1} R^{-2}_{{\rm MaNGA},i} - R^{-2}_{{\rm tpl},i}, \quad (7)$$

where $N_{\rm fit}$ is the number of fitted wavelength channels and $R_{{\rm tpl},i}$ and $R_{{\rm MaNGA},i}$ are, respectively, the spectral resolution ($R = \lambda/\Delta\lambda$) of the template library and MaNGA spectrum at the observed wavelength in channel i; i.e., the resolution vector of the template library is appropriately shifted to the best-fit redshift of the MaNGA spectrum for this calculation. In particular, we use the estimate of the resolution that *does not* include the integration of the LSF over the spectral channel, provided by the PREDISP extension in the DRP data cubes, because this is most appropriately matched to the pPXF method. Over the spectral region of the MILES-HC library, the MILES resolution is $\sim 16^{+7}_{-3}\%$ higher than the mean resolution of the spaxels in a MaNGA data cube ($R_{{\rm tpl},i} \sim 1.16\ R_{{\rm MaNGA},i}$), such that we should expect $\delta\sigma_{\rm inst} \sim 34$ km s^{-1}, according to Equation (7). Indeed, the median correction for the spaxels in a MaNGA data cube is $\delta\sigma_{\rm inst} \sim 32.6^{+9.0}_{-5.7}$ km s^{-1} (cf. Figure 17). We provide $\sigma_{\rm obs}$ from pPXF and $\delta\sigma_{\rm inst}$ from Equation (7) in the MAPS file. It is important to note that *we do not provide the corrected velocity dispersion*, σ_*. We leave it to the user to use Equation (6) to calculate σ_* for themselves. We discuss this decision in detail in Section 7.4.3 and provide some usage guidance in Section 7.7.

We describe the above calculation of $\delta\sigma_{\rm inst}$ as a first-order approximation because it makes the simplifying assumption that all spectral regions contribute equally to the determination of $\sigma_{\rm obs}$. However, we know that the influence of spectral features on the $\sigma_{\rm obs}$ measurements is roughly proportional to their equivalent width. That is, even with the spectral resolution vectors of the template and object spectra, an accurate measurement of $\delta\sigma_{\rm inst}$ is non-trivial because of the unknown relative influence of each spectral feature on the aggregate kinematics. Our recovery simulations, as presented in Section 7.5.2, have shown that our first-order estimate of $\delta\sigma_{\rm inst}$ from Equation (7) systematically underestimates the correction by a few percent. This leads to a systematic bias in σ_* of $\lesssim 5\%$ at $\sigma_* \approx 70$ km s^{-1}, with an increasing relative bias toward lower dispersion. However, our test of the accuracy of our first-order corrections presented toward the end of Section 7.4.3 shows no signs of a systematic error that is this large. We continue to improve the methodology used for the determination of the velocity dispersions and their corrections toward low dispersion, and will return to this topic in future papers.

7.2. Optimization and Performance Evaluation Data

The optimization and performance characterization of the DAP stellar kinematics are primarily based on three datasets:

1. A representative set of MaNGA spectra selected to span the full range in preliminary measurements of D4000 (Bruzual 1983), Hα equivalent width (EW), and stellar velocity dispersion (σ_*). All spectra are sorted into a three-dimensional grid defined by bin edges at

   $$\begin{aligned} {\rm D4000} &= \{1.2, 1.4, 1.5, 1.6, 1.8, 2.0, 2.2\}, \\ \sigma_{\rm obs} &= \{25, 50, 75, 100, 150, 200, 250\}, \text{ and} \\ {\rm H}\alpha\ {\rm EW} &= \{0, 2, 8, 16, 32\}, \end{aligned}$$

   with bins also for data below/above the first/last bin edge, for a total of 384 bins. When selecting spectra, any

[58] The value stored in the reference files is actually $cz_{\rm obs}$.

data cubes with the MANGA_DRP3QUAL bit set to CRITICAL by the DRP are ignored. We have selected the spectrum with the highest S/N_g in each bin (allowing for bins to have no relevant spectra) and then visually inspected the results to remove spectra with significant artifacts and with modeling failures; modeling failures in these cases are usually the result of interloping objects in the field outside the redshift boundary imposed by pPXF (± 2000 km s^{-1}). Our selection yielded 292 spectra for testing from 100 unique observations; 192 and 102 of these spectra have S/N_g larger than 30 and 60, respectively.

2. A representative set of four MaNGA data cubes for galaxies that span the mass-color ($\mathcal{M}_*$, NUV $-$ r) range of the MaNGA sample, listed in Table 2. Note that the two "blue" galaxies are the same as used by Belfiore et al. (2019); however, the two "red" galaxies chosen for Belfiore et al. (2019) were selected to also have noticeable emission lines, whereas the observations used here do not.
3. Fifty-six galaxies with multiple MaNGA observations provided in DR15 as listed in Table 1. As part of its ongoing quality control and calibration strategy, MaNGA has re-observed targets in fully identical plates (8256 and 8274), in identical IFUs on different plates (e.g., 7958-1901 and 9185-1901), and with different sized IFUs (e.g., 7960-12702 and 9185-3704). Repeat observations provide an ideal test-bed to empirically characterize the measurement uncertainties. We use these data to assess the robustness of our stellar kinematics here, our emission-line modeling results in Section 3 of Belfiore et al. (2019), and our spectral-index measurements in Section 10.3.

7.3. Template Library Comparison: MILES Versus MILES-HC

The motivation for the hierarchical-clustering analysis of the MILES library (Section 5) was to limit the templates used by pPXF to fit the MaNGA spectra, while not substantially affecting the fit quality or the resulting kinematics. Figure 12 shows the difference in the stellar kinematics as measured using the MILES and MILES-HC templates for the spectra in the first dataset described in Section 7.2. Compared to fits using the full MILES library, use of the MILES-HC templates when fitting the high-S/N MaNGA test spectra leads to an increase of the root-mean-square of the fractional residuals (fRMS) by typically 15% or less and of the reduced chi-square (χ^2_ν) by 30% or less. Figure 12 shows that this level of variation leads to marginal biases in the kinematics (see Emsellem et al. 2004, Figure B3). In detail, we find that the median difference and 68% and 95% difference invervals are intervals of $\Delta V = -1.3\ ^{+2.0\ +5.8}_{-1.8\ -6.3}$ and $\Delta\sigma_{\rm obs} = 0.9\ ^{+3.6\ +13.3}_{-4.1\ -17.7}$; the intervals are only moderately reduced if we restrict the test to the 102 spectra with $S/N_g > 60$. However, the difference in execution time is a factor of 25, as expected by the roughly $O(N_{\rm tpl})$ scaling of the pPXF method (i.e., the difference between fitting 985 MILES templates versus 42 MILES-HC templates). Given these two results (minimal effect on the kinematics and dramatically shortened computation time), we have used the MILES-HC library for both full-spectrum fitting algorithms in the DR15 execution of the DAP. The use of the MILES-HC library in the continuum modeling needed for the emission-line measurements has been discussed in detail by Belfiore et al. (2019).

Table 1
Repeat Observations

MaNGA ID	PLATEIFU (1)	(2)	(3)	(4)
1-113375	7815-9101	7972-12704	⋯	⋯
1-113379	7815-6101	7972-3701	⋯	⋯
1-113469	7815-12702	7972-12705	⋯	⋯
1-113525	7815-1902	8618-6103	7972-9102	⋯
1-113567	7815-12701	8618-1902	⋯	⋯
1-134760	8555-12701	8600-9102	⋯	⋯
1-137801	8247-3702	8249-3701	⋯	⋯
1-137845	8250-9101	8249-6104	⋯	⋯
1-137853	8250-3702	8249-12705	⋯	⋯
1-149686	8997-12701	8996-12705	⋯	⋯
1-166754	8459-3704	8461-12703	⋯	⋯
1-166919	8459-3702	8461-3704	⋯	⋯
1-166932	8459-3701	8461-6104	⋯	⋯
1-167356	8456-6104	8454-6103	⋯	⋯
1-177236	7958-1901	9185-1901	⋯	⋯
1-177250	7958-3703	9185-3702	⋯	⋯
1-178442	7962-6101	9085-3703	⋯	⋯
1-178443	7962-6104	9085-3704	⋯	⋯
1-178473	7962-3701	9085-3701	⋯	⋯
1-209770	9031-6102	9036-6104	⋯	⋯
1-209772	9031-3704	9036-3703	⋯	⋯
1-209786	9031-3701	9036-1901	⋯	⋯
1-209823	9031-12701	9036-12703	⋯	⋯
1-210186	9031-12705	9036-6101	⋯	⋯
1-210604	8600-3702	8979-12704	⋯	⋯
1-210611	8600-1902	8979-3703	⋯	⋯
1-210700	8603-12701	8588-3701	⋯	⋯
1-211017	8312-12703	8550-9102	⋯	⋯
1-235398	8326-12701	8325-12703	⋯	⋯
1-235530	8329-1901	8326-3701	⋯	⋯
1-255691	8256-6102	8274-6102	8451-3702	⋯
1-255959	8256-9102	8274-9102	⋯	⋯
1-256048	8256-6103	8274-6103	8451-6102	⋯
1-256104	8256-9101	8274-9101	8451-9101	⋯
1-256456	8256-12703	8274-12703	8451-12704	⋯
1-256457	8256-1902	8274-1902	8451-1902	⋯
1-258311	8261-1901	8262-1901	⋯	⋯
1-266074	8329-3703	8333-12704	⋯	⋯
1-277103	8256-6101	8274-6101	8451-6101	⋯
1-277154	8256-1901	8274-1901	⋯	⋯
1-277159	8256-3702	8274-3702	⋯	⋯
1-277161	8256-3701	8274-3701	⋯	⋯
1-277162	8256-12702	8274-12702	⋯	⋯
1-277691	8256-12701	8274-12701	⋯	⋯
1-277858	8256-3703	8274-3703	8451-3703	⋯
1-278485	8256-3704	8274-3704	8451-3704	⋯
1-456757	8479-3703	8480-3701	8953-3702	9051-6103
1-548221	8567-12702	8239-6104	⋯	⋯
1-558910	8256-6104	8274-6104	8451-6103	⋯
1-558912	8256-12704	8274-12704	8451-12701	⋯
1-561017	7960-12702	9185-3704	⋯	⋯
1-569225	8329-3701	8333-12701	⋯	⋯
1-587938	8256-12705	8274-12705	8451-12702	⋯
1-592881	8329-3704	8333-3702	⋯	⋯
1-635503	7815-12705	8618-6101	⋯	⋯
1-93876	8555-3704	8484-9101	⋯	⋯

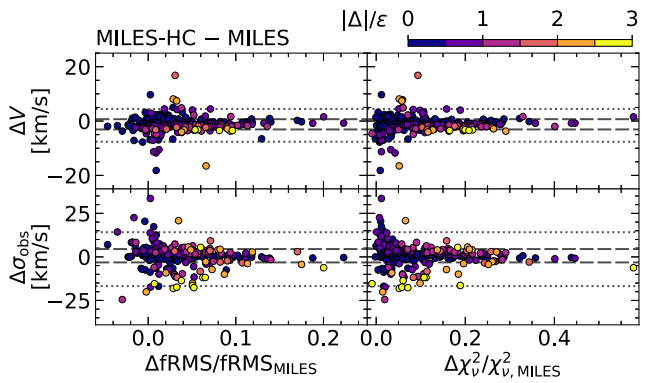

Figure 12. Difference between the stellar velocity (top), V, and the "observed" stellar velocity dispersion (bottom; see Section 7.1.5), $\sigma_{\rm obs}$, measured for the first dataset listed in Section 7.2 using our implementation of pPXF in the DAP and either the full MILES or MILES-HC template library. Differences are plotted as a function of the fractional change in the root-mean-square of the fractional residuals (fRMS; left) and the reduced chi-square (χ^2_ν; right). Positive values mean the MILES-HC metric or measurement is larger. The difference in the measurement relative to its error, $|\Delta|/\epsilon$, is represented by the point color, according to the color bar. The horizontal lines show the interval enclosing 68% (dashed) and 95% (dotted) of all data.

Table 2
Representative Observations

MaNGA ID	PLATEIFU	M_i	$\log(\mathcal{M}_*/\mathcal{M}_\odot)$	NUV − r
1-113379	7815-6101	−17.5	8.7	2.2
1-339041	8138-12704	−22.1	11.0	3.7
1-377176	8131-3702	−20.6	10.3	5.5
1-113379	8131-6102	−22.3	11.1	5.3

7.4. Design Choices

We explore and justify three core design choices implemented by our fitting algorithm, as described in Section 7.1. Namely, we quantify the effect of the algorithmic down-selection of the templates on the resulting kinematics in Section 7.4.1, we describe tests performed to optimize the order of the additive Legendre polynomial used in all fits in Section 7.4.2, and we justify our use of the templates at their native resolution in Section 7.4.3.

7.4.1. Algorithmic Down-selection of Templates

In the same vein of reducing its execution time while minimizing the degradation of the fit quality and stellar kinematics, the DAP algorithmically down-selects templates from the larger MILES-HC library by tuning the templates used to fit the individual spectra of each galaxy. This is done by first fitting the global spectrum and only using those templates with non-zero weights in the subsequent fits (Section 7.1.3). This approach reduces the per-data-cube execution time of the stellar-kinematics module by a factor of 2.5–3.[59] How does this further down-selection of the templates change the resulting stellar kinematics when compared to a fit that always uses the full MILES-HC library?

[59] Use of this algorithmic down-selection with the full MILES library leads to execution times that are a factor of ∼2 slower than when used with the MILES-HC library. Of all the tests that we have performed, we have consistently found that the roughly $O(N_{\rm tpl})$ execution-time scaling of the pPXF algorithm holds. In this example, the number of non-zero templates selected from the global fit using the full MILES library is approximately twice as many as selected when using the smaller MILES-HC library.

Figure 13 compares V and $\sigma_{\rm obs}$ (Equation (6)) measurements for three different fit approaches: ALL—a fit that uses all templates from the MILES-HC library for all spectra (i.e., no down-selection is performed); NZT—a fit that only uses the templates given a non-zero weight in a fit to the global spectrum; and GLB—a fit that only uses a *single* template constructed using the weights determined for the fit to the global spectrum. These fits were performed using the four example data cubes discussed in point 2 of Section 7.2 without any spatial binning and limited to spaxels with $S/N_g > 1$.

The top row of Figure 13 shows S/N_g versus $\sigma_{\rm obs}$ resulting from the nominal DR15 approach (NZT), primarily as a reference for the differences seen in the following panel rows. The next two panel rows show the difference in V and $\sigma_{\rm obs}$ between the ALL and NZT approaches. As given by the point color, the majority of the differences between the measurements are below 50% of the error. The main exceptions to this are the $\Delta\sigma_{\rm obs}$ measurements near the center of the massive blue galaxy. This galaxy has the strongest broad-band color gradient from its center to its outskirts, so it is reasonable to find that the templates selected by a fit to the global spectrum may not capture the templates relevant to relatively small and/or low-surface-brightness regions. Indeed, this is much more apparent when only a single template determined by the fit to the global spectrum is used, as is the case in the bottom two rows of Figure 13, which compares the GLB and NZT approaches. Here, the differences in both V and $\sigma_{\rm obs}$ for the high-mass blue and high-mass red galaxies can be more than 3 times the measurement error near their centers. For this reason, we chose not to adopt the GLB approach.

However, the NZT and ALL approaches are reasonably consistent. In all cases, ΔV is small compared to its error. For the few spaxels where $\Delta\sigma_{\rm obs}$ is up to 1.5 times its error, the differences are small in both an absolute sense (2–4 km s^{-1}) and a relative sense (1%–2%). Tests of the relative biases in the NZT approach are ongoing; however, given the factor of 2.5–3 decrease in execution time and acceptable level of change to the kinematics from these few examples, we have adopted the NZT approach for the analysis of all data cubes in DR15.

7.4.2. Low-order Polynomial Manipulation of the Continuum

The low-order polynomials included in pPXF model spectra aim to compensate for any subtle mismatch between the spectral templates and the science spectra that change slowly with wavelength. A mismatch may arise due to, e.g., flux-calibration errors, internal or foreground reddening, and even library incompleteness. The order of the polynomial should be sufficiently low to avoid impacting individual line fits, except insofar as additive polynomials can modulate the line equivalent width (but not shape).

However, given that polynomials allow for templates to be selected that can fit detailed lines while relaxing mismatch in their continuum shape, the inclusion of these functions allows for different solutions. These solutions can, in principle, differ at the level of individual line-profile fits, examples of which are seen clearly in Belfiore et al. (2019). Although the focus of our analysis in this companion paper is on the strong H and He lines in the blue for young stellar populations, Figure 14 therein shows that differences extend more broadly across the spectrum and, by inference therefore, to the weak metal lines that influence the kinematics solution of the pPXF fits.

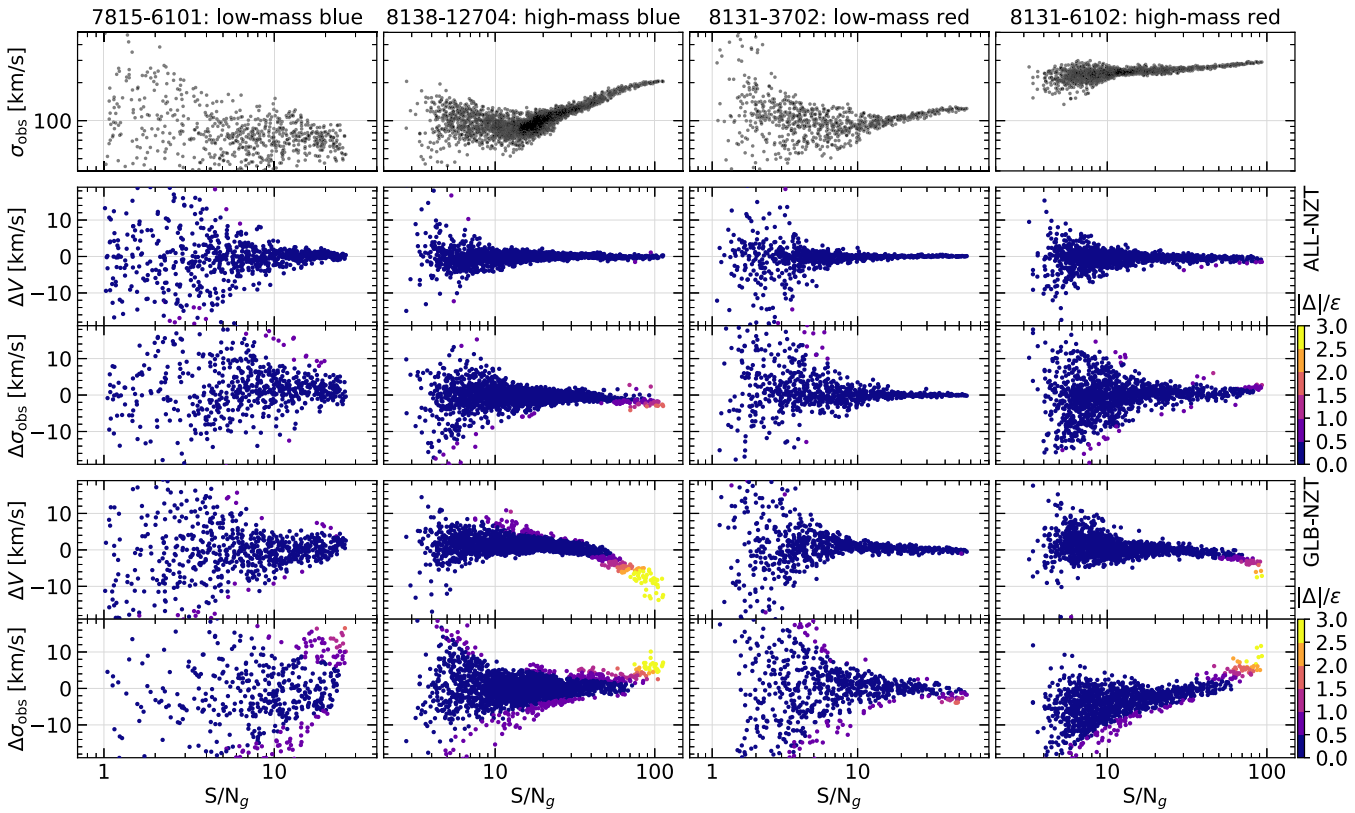

Figure 13. Relevance of the template down-selection approach to the stellar kinematics in four representative MaNGA galaxies (Table 2; Section 7.2) as labeled at the top of each panel column. The top row shows $\sigma_{\rm obs}$ as a function of S/N$_g$ for each galaxy resulting from the NZT approach (see Section 7.4.1). The second and third panel rows show, respectively, the velocity difference, ΔV, and velocity-dispersion difference, $\Delta\sigma_{\rm obs}$, between the ALL and NZT approaches. The absolute value of the difference relative to its error, $|\Delta|/\epsilon$, is given by the point color. The fourth and fifth panel rows show the differences when comparing the GLB and NZT approaches. The difference between the GLB and NZT approaches are quite different for galaxies with significant stellar-population (color) gradients, like 8138-12704, whereas the ALL and NZT approaches are more consistent.

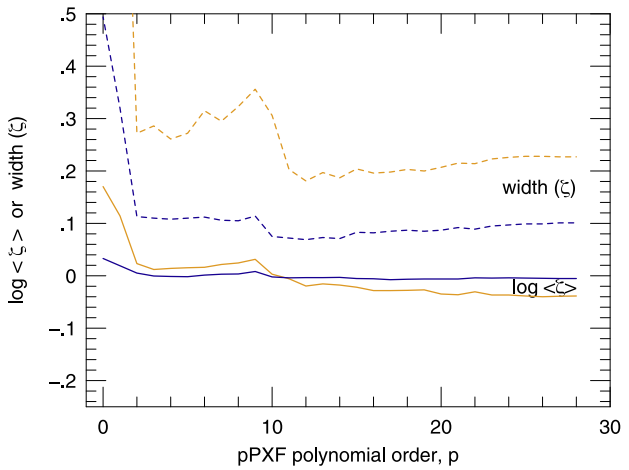

Figure 14. Assessment of the influence of the order (p) of the additive Legendre polynomial used during the stellar-kinematics fit on the best-fitting $\sigma_{\rm obs}$. Statistics of a representative sample of MaNGA spectra (Section 7.2) are for the normalized velocity dispersion (ζ), and they include the mean and standard deviation (dark-solid and dashed lines, respectively) and the median and median absolute deviation (light-solid and dashed lines, respectively). The normalization is the mean value of $\sigma_{\rm obs}$ for each spectrum over all polynomial orders, p; i.e., $\zeta = \sigma_{\rm obs}/\langle\sigma_{\rm obs}\rangle_p$. Note the standard deviation is $\sim$1.5 times larger than the median absolute deviation for a normal distribution.

Our concern here is whether the inclusion of polynomials in the pPXF fitting alters the values of the derived stellar kinematics in a quantitatively significant fashion. We are primarily concerned here with velocity dispersion because these (and higher) moments are believed to be more likely affected by detailed template changes; i.e., the infamous "template mismatch" (e.g., Figure 6 of van der Marel et al. 1994).

To test for systematic differences in the derived stellar velocity dispersion, we used the 292 representative spectra that span a range of parameters (Section 7.2). We used pPXF to fit each spectrum with the MILES-HC template library but with an additive Legendre polynomial of order p ranging from 0 to 30 ($p = 0$ represents the case for no polynomial term). Of the 292 spectra, 290 were successfully fit for all orders. For *each* of these 290 spectra, we determined the mean $\sigma_{\rm obs}$ over all polynomial orders for the individual spectrum, $\langle\sigma_{\rm obs}\rangle_p$, and considered the statistical trend of $\zeta = \sigma_{\rm obs}/\langle\sigma_{\rm obs}\rangle_p$ with polynomial order.

The mean, median, standard deviation, and median absolute deviation of ζ for the full sample are shown in Figure 14. Strong trends exist in the characteristic value and scatter of ζ for orders $p < 3$. There is some evidence for other trends when $p \geqslant 9$, but this is less evident using robust statistics. Further exploration reveals that the strength of the trends in the mean and standard deviation at larger polynomial orders correlate with *decreasing* S/N. The depth of the MaNGA survey is uniform enough that S/N is well correlated with surface-brightness over all observations. Because of this and the correlation of surface-brightness with D4000, Hα EW and σ_* (younger stellar populations have small σ_* and tend to be

found in the outskirts of galaxies where the surface-brightness is low), the strength of the statistical trends at high polynomial order correlates broadly with many variables describing the spectral sample. That said, the systematic changes in σ_* with changing polynomial order is small (below a few percent) for orders $p > 3$, and σ_* for values $p < 9$ are relatively immune to systematics at lower S/N. A decision to use $p = 8$ was made early in the development of the DAP, and this more detailed analysis demonstrates that there is no compelling reason to revisit that choice.

7.4.3. Spectral-resolution Matching

The common approach to measuring stellar kinematics is to use template spectra that have a spectral resolution—or instrumental dispersion, $\sigma_{\rm inst}$—that is matched to the galaxy data. In fact, when fitting higher moments of the LOSVD with pPXF, the definition of the Gauss–Hermite parameterization (van der Marel & Franx 1993) requires this to be the case. However, under the simplistic assumption of $\sigma_{\rm inst} \approx 70$ km s^{-1} for MaNGA, we find that 40% of all DAP-analyzed spectra in DR15 have $\sigma_* < \sigma_{\rm inst}$, and half of all MaNGA data cubes show $\sigma_* < \sigma_{\rm inst}$ for at least 32% of their spectra. Thus, it will be difficult, if not impossible, to reliably measure the higher-order moments of the LOSVD (see Cappellari & Emsellem 2004) for many of the spectra in DR15. For this reason, we have not measured the Gauss–Hermite moments in this data release but have instead restricted our measurements to the first two velocity moments alone (V and σ_*). Future improvements of the DAP may include fits of the higher-order moments for a relevant subset of MaNGA spectra.

Free from the *requirement* of resolution matching the input spectra, we explore the difference between fits performed with and without the matched-resolution spectra.

First, in the limit of a Gaussian LOSVD and Gaussian LSFs in both spectra, we note that there is no mathematical difference between first convolving the template spectrum with a Gaussian kernel based on the resolution difference $\delta\sigma_{\rm inst}$ before executing the pPXF fit versus subtracting that difference in quadrature from the pPXF result, as in Equation (7).

Second, because of the Doppler shift between the template and galaxy spectra, an offset between the spectral resolution of the template and galaxy data is effectively inevitable, even for spectra observed with identical instrumental resolution. This intrinsic difference means that, at some level, all velocity-dispersion measurements require a correction for the detailed, wavelength-dependent difference of $\sigma_{\rm inst}$ between the template and galaxy spectra (see Westfall et al. 2011).

Third, we therefore expect $\delta\sigma_{\rm inst} > 0$ and, given that measurements of $\sigma_{\rm inst}$ will always have some uncertainty, it is useful to understand how those uncertainties for both the galaxy and template spectra propagate to the uncertainty in $\sigma_{\rm obs}$. We explore this in Appendix B and show that, under some nominal assumptions, it is advantageous in terms of the error budget to use template spectra with higher spectral resolution than the galaxy data. Although the resolution difference between MaNGA and MILES is modest (the median $\sigma_{\rm inst}$ for MaNGA is ∼16% larger than for MILES over their common spectral range), our calculations expect a substantial improvement in the $\sigma_{\rm obs}$ uncertainty, particularly as σ_* becomes less than the MaNGA $\sigma_{\rm inst}$.

Fourth, as $\sigma_{\rm obs}$ approaches 0, its error distribution will become more significantly non-Gaussian, which complicates the handling of the data both in terms of aggregation and model fitting. In the limit where measurements of $\sigma_{\rm obs}$ are similar to fitting a Gaussian function to a set of data, we should expect its posterior probability to follow an inverse-gamma distribution (MacKay 2003, Section 24.1). When the ratio of its mean to its standard deviation is large, the inverse-gamma distribution is well-approximated by a Gaussian; however, as this ratio decreases, the inverse-gamma distribution exhibits increasingly significant positive skew. Thus, in addition to gains in terms of the error budget (Appendix B), one benefits from having $\delta\sigma_{\rm inst} > 0$ in terms of the form of the $\sigma_{\rm obs}$ error distribution: as σ_* approaches 0, $\sigma_{\rm obs}$ approaches the constant $\delta\sigma_{\rm inst}$, which can be used to limit the ratio of the expectation value of $\sigma_{\rm obs}$ to its variance at fixed S/N. Again, the difference in spectral resolution between MaNGA and MILES is rather modest; however, fitting the MaNGA spectra with the MILES templates at their native resolution enables us to mitigate some of the issues with the error distribution of $\sigma_{\rm obs}$ at low σ_*. In particular, this includes reducing the number of measurements that hit the $\sigma_{\rm obs} \approx 0$ boundary imposed by pPXF.[60]

It is the latter consideration, as motivated later on, that was the main driver of our decision to perform the fit of the first two moments of the stellar LOSVD without matching the resolution of the templates to that of the galaxy, instead keeping them at their native resolution.

Before continuing, we note that the public Python version of pPXF employed by the DAP uses a novel trust-region implementation of the Levenberg–Marquardt (see Section 10.3 of Nocedal & Wright 2006) least-squares nonlinear optimization algorithm, which rigorously deals with bound constraints or fixed and tied parameters.[61] A best-fit parameter at the boundary of the allowed region does not necessarily indicate a convergence failure, only that the minimum χ^2 is at the boundary.

For illustration purposes, we fit the individual spaxels for ∼100 data cubes drawn from the set of repeat observations listed in Table 1. Each spectrum is fitted twice, once with the MILES-HC library at its native resolution and again after matching the template data to the MaNGA spectral resolution (Appendix A). We bin the results by their S/N and σ_*—i.e., the direct output from pPXF for the matched-resolution case and after applying the velocity-dispersion corrections for the resolution-offset case (Section 7.1.5). The top row of Figure 15 shows the total number of fitted spaxels, $N_{\rm fit}$, in each fitting mode. Any difference between the top two panels of Figure 15 is due to spaxels being located in different σ_* bins, which are mostly seen for the lowest σ_* bin. The bottom two panels of Figure 15 compare $N_{\rm fit}$ to the number of spaxels that are within the pPXF bounds on $\sigma_{\rm obs}$, $N_{\rm bound}$.

There is a clear difference in the number of fits with $\sigma_{\rm obs}$ within the trust-region boundary that result from the treatment of the spectral-resolution difference, particularly at low σ_* and low S/N. At virtually infinite S/N and with zero template or LSF mismatch, pPXF can measure $\sigma_{\rm obs}$ well below the instrumental dispersion, as shown by Cappellari (2017, Figure 2), meaning that this discrepancy is not an intrinsic issue with the pPXF algorithm itself. Instead, this is because

[60] To be precise, by default $\delta_v/100 < \sigma_{\rm obs} < 1000$ km s^{-1}, where $\delta_v = 10^{-4}$ ln(10) $c = 69$ km s^{-1} is the velocity step per spectral sample of the DRP log-linear binned data cubes.

[61] The Python version of pPXF has used method='capfit' in place of MPFIT (Markwardt 2009) since version 6.5.

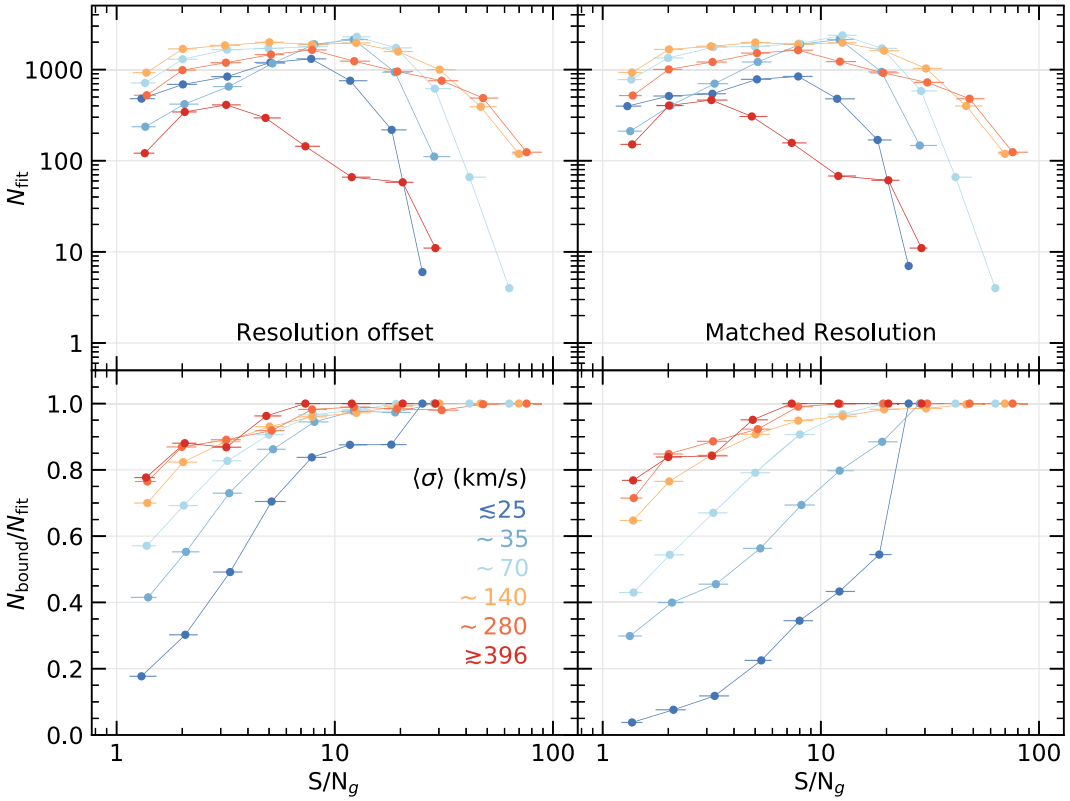

Figure 15. For individual spaxels from ∼100 data cubes drawn from Table 1, we show the total number of fitted spectra ($N_{\rm fit}$; top) and its ratio compared to the number of fitted spectra with $\sigma_{\rm obs}$ within the pPXF trust-region boundary ($N_{\rm bound}$; bottom). The results are shown for fits using template spectra that have (right) and have not (left) had their resolution matched to the MaNGA resolution (Appendix A). The results are binned by S/N$_g$ and the *astrophysical* velocity dispersion, σ_*; the latter are binned geometrically with the color and geometric center at the value given by the legend; e.g., the 70 km s^{-1} bin includes measurements between $50 \lesssim \sigma_* \lesssim 100$ km s^{-1}. The bottom panels demonstrate that the number of fitted spectra with viable σ_* measurements from pPXF is dramatically larger when a resolution offset exists between the data and the templates, which allows for improved assessments of the error distribution for low σ_*.

spectral-resolution uncertainties—even at the very modest level expected of either MILES (∼2%; Beifiori et al. 2011; Falcón-Barroso et al. 2011) or MaNGA (∼3%)—and noise can lead to matched-resolution templates with broader lines than the galaxy spectra and drive the pPXF fit to the $\sigma_{\rm obs} \approx 0$ boundary. The pedestal offset in the spectral resolution of the MaNGA and MILES-HC spectra helps avoid this and yields measurements that more often have non-zero $\sigma_{\rm obs}$, even with the modest spectral resolution difference between MaNGA and MILES.

To account for the effect of the resolution offset on the measured $\sigma_{\rm obs}$ reported by pPXF, we must calculate a correction that removes this difference and provides the astrophysical velocity dispersion of the stars, σ_* (Section 7.1.5; Equation (6)). We assess the accuracy of the velocity-dispersion corrections provided in DR15 in the following two ways.

First, we compare σ_* determined with and without matching the template resolution to the MaNGA data; we exclude any measurements from the latter with $\sigma_{\rm obs} \leqslant \delta\sigma_{\rm inst}$. Figure 16 illustrates that the measurements made using either method are consistent to within 1%–2% for $\sigma_* \gtrsim 50$ km s^{-1}. Systematic differences become more significant as σ_* becomes small relative to $\sigma_{\rm inst} \sim 70$ km s^{-1}. There are two considerations that lead to this discrepancy. First, our current approach (Section 7.1.5) may produce slight underestimates of $\delta\sigma_{\rm inst}$ that only yield significant differences at low $\sigma_{\rm obs}$. This underestimation is qualitatively consistent with results from our idealized simulations and may motivate a change to the determination of the correction used in future data releases. Second, at low σ_*, more measurements from the resolution-

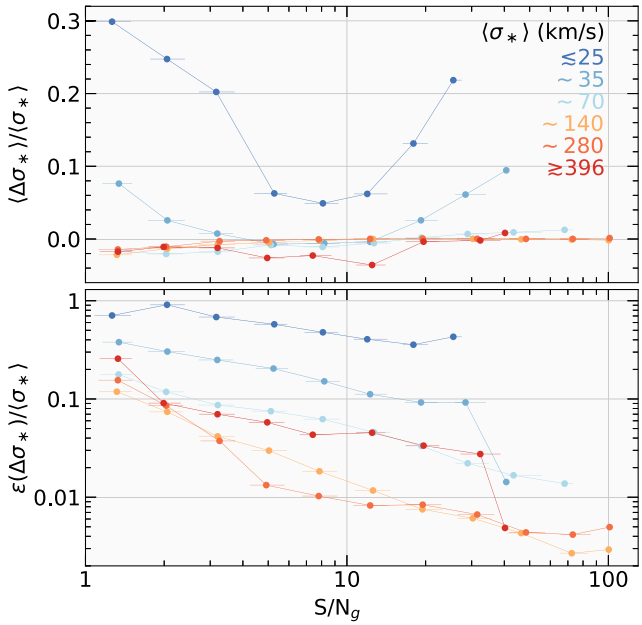

Figure 16. The difference between the astrophysical dispersion, σ_*, obtained for spectra analyzed using templates that have and have not had their spectral resolution matched to the MaNGA data (see Section 7.4.3). Data are binned similarly to Figure 15. The top panel shows the mean percentage difference and the bottom panel shows the percent scatter in the difference; both are plotted as a function of S/N$_g$. The difference between the two measurements, $\Delta\sigma_*$, is positive if the measurement made with mismatched resolution is larger than the measurement made with matched resolution.

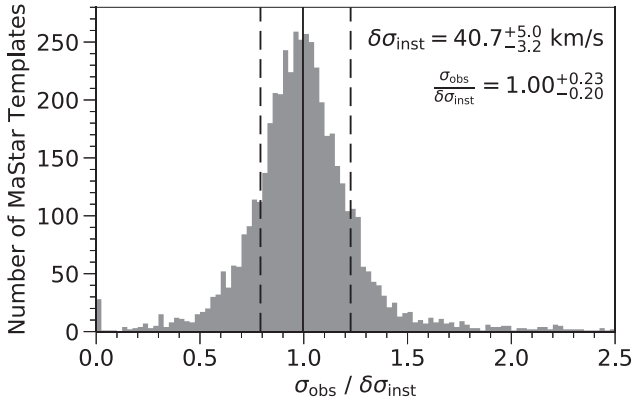

Figure 17. First-order velocity-dispersion correction based on the resolution vectors of a set of ∼5000 MaStar stellar spectra and the MILES-HC template library, $\delta\sigma_{\rm inst}$, are compared with the effective velocity dispersions from pPXF, $\sigma_{\rm obs}$, determined by fitting MILES-HC to the MaStar spectra. The plotted histogram of the ratio of these two measurements has a median of unity with an inner-68% interval of ±20%, as shown in the upper-right corner of the plot. The median $\delta\sigma_{\rm inst}$ and its inner-68% interval are also given.

offset approach have $\sigma_{\rm obs} > \delta\sigma_{\rm inst}$ than there are measurements of $\sigma_{\rm obs}$ within the trust-region boundary in the matched-resolution approach. This leads to more measurements with systematically higher σ_* in the resolution-offset case.

Second, we fit a subset of 5000 randomly selected stellar spectra from the MaStar empirical stellar library (Yan et al. 2019) using a similar setup to that used by the DAP. Assuming negligible stellar rotation or other atmospheric broadening effects, measurements of $\sigma_{\rm obs}$ determined by fitting the MILES-HC templates to the MaStar spectra provide direct measurements of $\delta\sigma_{\rm inst}$ that we can compare to our first-order calculation (Section 7.1.5). Figure 17 shows that the measurements of $\sigma_{\rm obs}$ and $\delta\sigma_{\rm inst}$ are very consistent: the median $\delta\sigma_{\rm inst}$ for the MaStar spectra is $40.7^{+5.0}_{-3.2}$ km s^{-1}, and we find that these measurements are consistent with the directly measured values, $\sigma_{\rm obs}$, with less than 1% difference in the median and an inner-68% interval of approximately ±20%.

Finally, we note that the number of measurements with $\sigma_{\rm obs} \leqslant \delta\sigma_{\rm inst}$ for the resolution-offset measurement is nearly the same as the number of measurements with $\sigma_{\rm obs} = 0$ for the matched-resolution fits, except at the very lowest σ_*. One expects this given the consistency of the comparison shown in Figure 16. Therefore, simply ignoring measurements with $\sigma_{\rm obs} \leqslant \delta\sigma_{\rm inst}$ is one possible approach we discuss in Section 7.7 in terms of how to use the DR15 data, and this should lead to results that are consistent with matched-resolution fits that ignore measurements at the lower trust-region boundary of $\sigma_{\rm obs}$. However, one should consider alternative approaches presented in Section 7.7 that take advantage of the resolution offset and help mitigate σ_* biases at low σ_* and low S/N.

7.5. Performance

7.5.1. Empirical Uncertainties from Repeat Observations

We use the first two observations of the 56 galaxies with multiple MaNGA observations (Table 1) to characterize the trends of the errors in our stellar kinematics with S/N$_g$ and test the accuracy of the formal error estimates returned by pPXF (as described toward the end of this Section). The data-cube-reconstruction algorithm generally makes a comparison of multiple observations of a single target straightforward, even for IFUs of different size. The target is typically centered in the data cube, such that registering the WCS coordinates of MaNGA data from different IFUs is a simple offset of the arrays to align their spatial center. In the few cases where this is not true, we align the data by simply interpolating the relevant values to a common coordinate grid.

Figure 18 compares the two observations of MaNGA galaxy 1-149686 (8996-12705 and 8997-12705) as an example of the differences in the stellar kinematics that we quantify statistically across all repeat observations in Figure 19. No registration was required for the observations in Figure 18 because they have identical array sizes and WCS coordinates. We show the differences in the g-band surface-brightness (I_g, calculated directly for each spaxel), the stellar velocity (V), and the corrected stellar velocity dispersion (σ_*); for the latter, we show maps of σ_*^2 that include negative values resulting from Equation (6) and maps of σ_* where those spaxels with $\sigma_*^2 < 0$ have been removed.

Figure 18 shows general consistency between the mapped quantities and leads one to a sense of the influence of random errors on the DRP and DAP products. First, note that the kinematic residuals do not generally vary on a spaxel-by-spaxel scale, but more beam-by-beam. This is a natural effect of the significant spatial covariance (Figure 7) in the MaNGA data (see Section 2). Second, in the lowest surface-brightness pixels, the residuals in σ_* begin to vary spaxel-by-spaxel, which is driven by systematic error in these measurements at low S/N and low σ_*.

To empirically assess the random errors in the DAP stellar kinematics in DR15, we similarly register and perform a spaxel-by-spaxel comparison of the measurements from the first two observations of all 56 galaxies with multiple MaNGA observations in Table 1. Here and henceforth, we only focus on our standard approach for DR15, which is to measure the stellar kinematics *without* resolution-matching the template and galaxy spectra and after correcting the pPXF output $\sigma_{\rm obs}$ to obtain σ_*. Spaxel pairs at identical spatial locations in the repeat observations are binned (geometrically) as a function of S/N$_g$ and σ_*. Pairs with a measurement of $\sigma_*^2 < 0$ for either observation are excluded. For each S/N$_g$–σ_* bin, we calculate the mean reduced chi-square (χ^2_ν) of the fit, the mean σ_* ($\langle\sigma_*\rangle$), the mean and standard deviation in the difference in the kinematics (e.g., $\langle\Delta V\rangle$ and $\delta(\Delta V)$), and the mean of the DAP-reported error in the difference (e.g., $\langle\epsilon[\Delta V]\rangle$). The results of these calculations are shown in Figure 19.

From the top-middle panel of Figure 19, we note the increase in χ^2_ν toward large S/N$_g$. This is an expected trend, due to the systematic differences in the MILES-HC library and the observed galaxy spectra becoming a more significant fraction of the random errors in the flux density. These results are consistent with our findings for the full DR15 sample, as discussed more at length as part of our general performance assessments of the DAP in Section 11.

The bottom two rows of Figure 19 are meant to assess, from left to right panels, the systematic error, the random error, and the accuracy of the DAP-reported errors in V and σ_*. The left panels show that, in the mean, there is little systematic difference between the V and σ_* measurements relative to the scatter in the difference (middle panels). These measurements are very useful to assess the repeatability of the measurements and the reliability of the DAP formal errors in a relative sense.

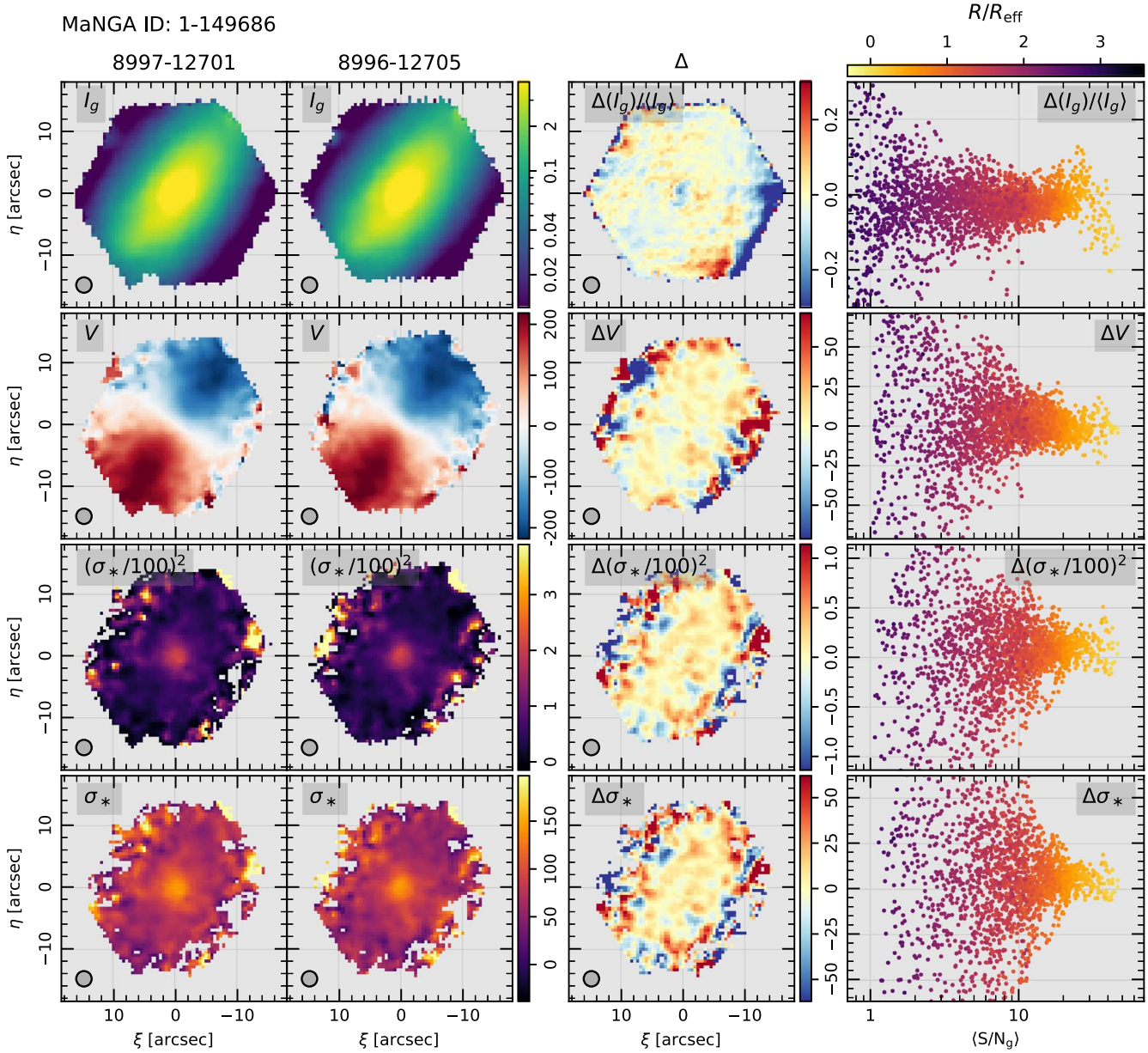

Figure 18. The difference in the stellar kinematics of MaNGA galaxy 1-149686 as measured by the DAP using observations 8996-12705 and 8997-12705. From top to bottom, we plot the mean g-band surface-brightness measured from the MaNGA data cubes, the stellar velocity in km s^{-1}, the corrected stellar velocity dispersion squared (allowing for negative values) in (km s^{-1})^2, and the corrected stellar velocity dispersion, σ_*, in km s^{-1}. The two columns to the left show the measurements from each observation, the third column shows their difference, and the right column shows the differences as a function of S/N$_g$, colored by semimajor-axis radius.

However, these data cannot assess the systematic error in an absolute sense because we do not know the *true* kinematics and are instead comparing two measurements that may both suffer from an absolute systematic error. Moreover, by binning the data by the *measured* σ_*, the statistics we have calculated are more of an assessment of the distribution of the data within the bin instead of an assessment of the distribution of data about the intrinsic σ_*.

The expectation found by many authors (e.g., Jorgensen et al. 1995; Westfall et al. 2011) is that random errors in both V and σ_* should be directly proportional to σ_*—at fixed S/N, the centroid uncertainty increases with the line width—and inversely proportional to spectral S/N—higher S/N spectra provide smaller random errors. The middle column of the bottom two panel rows in Figure 19 demonstrates that this is generally true for the DAP results, with some notable exceptions. For measurements with $\sigma_* \gtrsim 100$ km s^{-1} (i.e., those measurements following the red and orange lines in Figure 19), we find that the velocity errors are roughly $\delta(\Delta V) \approx \langle\sigma_*\rangle(\mathrm{S/N}_g)^{-1}$. That is, the velocity errors are 10% of σ_* at S/N$_g$ = 10. The errors in σ_* are slightly larger than that but are also well-approximated by a single proportionality constant for $\sigma_* \gtrsim 100$ km s^{-1}. However, for $\sigma_* \lesssim 100$ km s^{-1}, when $\sigma_* \sim \sigma_{\rm inst} \sim 70$ km s^{-1}, the proportionality constant changes such that the errors become a more substantial fraction of σ_*. In the lowest $\langle\sigma_*\rangle$ bin, the V and σ_* errors are increased to, respectively, ∼30% and ∼60% of σ_* at S/N = 10. We expect this is because the width of the observed features in the spectrum

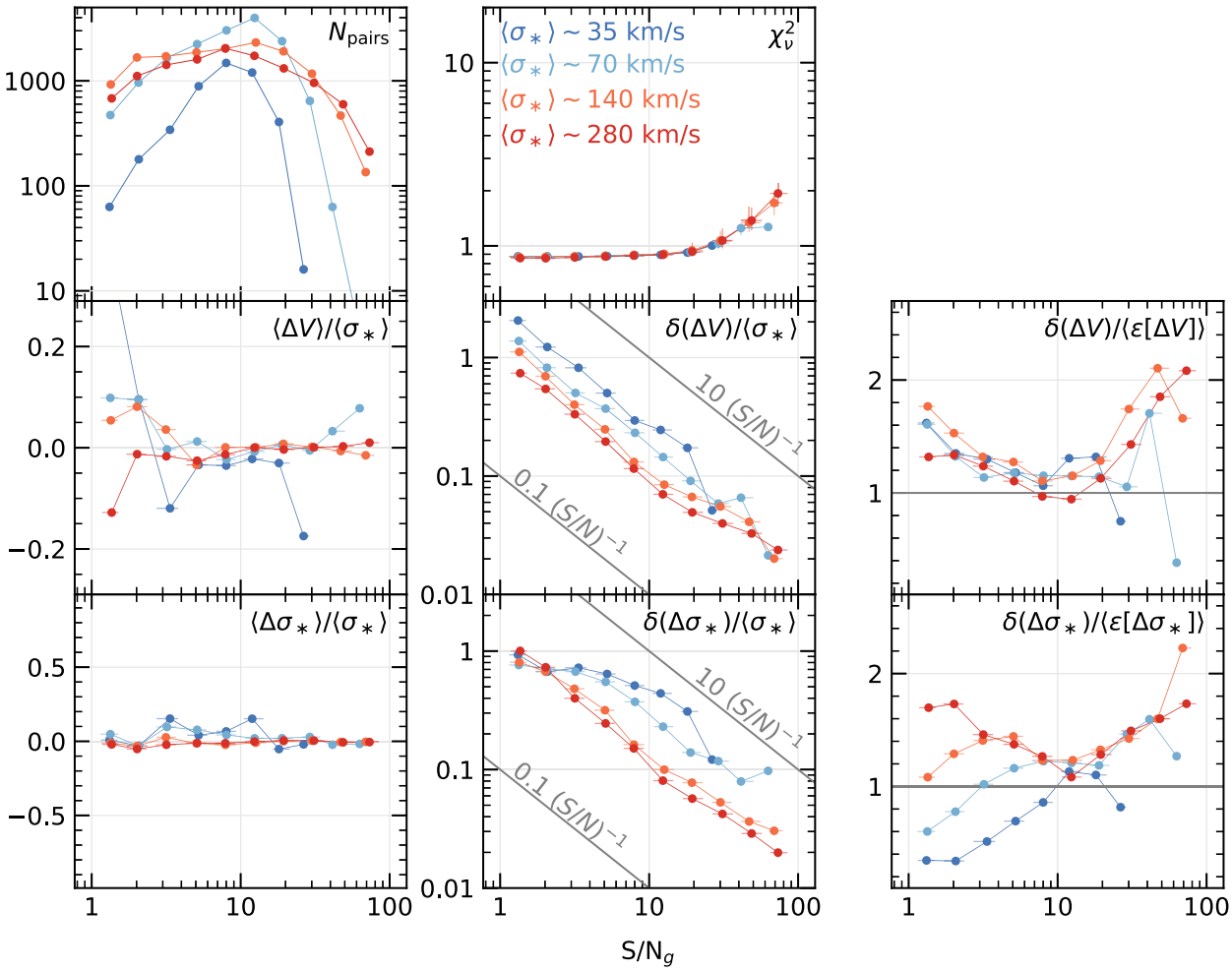

Figure 19. A statistical comparison of the spaxel-by-spaxel measurements of the stellar kinematics from repeat observations of a set of 56 galaxies (Table 1). Results are binned as a function of S/N$_g$, according to the abscissa of each panel, and corrected velocity dispersion (σ_*), according to the color in the legend. From left-to-right, the top row shows the number of spaxel pairs compared and the mean χ^2_ν for all spaxels in each S/N$_g$–σ_* bin. The middle and bottom rows, respectively, assess the robustness of the velocity, V, and velocity dispersion, σ_*, and their DAP-reported errors (ϵ). For these two rows: the left panels shows the mean difference between the two kinematic measurements—$\langle\Delta V\rangle$, $\langle\Delta\sigma_*\rangle$—relative to the mean σ_*, an assessment of systematic error. The middle panels show the standard deviation in the difference between the two kinematic measurements—$\delta(\Delta V)$, $\delta(\Delta\sigma_*)$—relative to the mean σ_*, an assessment of random error. The right panels compare the random error in the repeat observations to the mean formal error for each kinematic measurement differences—$\langle\epsilon[\Delta V]\rangle$, $\langle\epsilon[\Delta\sigma_*]\rangle$, an assessment of the accuracy of the formal errors.

become increasingly dominated by $\sigma_{\rm inst}$ such that the kinematics errors become increasingly independent of σ_*. Finally, although $\delta(\Delta V) \propto \langle\sigma_*\rangle({\rm S/N}_g)^{-1}$ holds for all σ_* bins at all S/N$_g$, the proportionality is lost for $\delta(\Delta\sigma_*)$ at low σ_* and low S/N$_g$, where $\delta(\Delta\sigma_*)$ becomes roughly independent of S/N$_g$. We expect this happens because a larger number of measurements hit the trust-region boundary on $\sigma_{\rm obs}$ in this regime, due to the large uncertainties.

The right panels in Figure 19 compare the uncertainties in the kinematics estimated directly from the repeat observations, $\delta(\Delta V)$ and $\delta(\Delta\sigma_*)$, to the mean of the errors provided by the formal calculation in pPXF, $\langle\epsilon[\Delta V]\rangle$ and $\langle\epsilon[\Delta\sigma_*]\rangle$. Formal uncertainties provided directly by pPXF are based on the covariance matrix of the fitted parameters, which can be computed using the Hessian matrix generated as a byproduct of the Levenberg–Marquardt nonlinear least-squares optimization algorithm; see Section 15.5 of Press et al. (2007).[62] This is a standard approach to measurement errors made under the (strong) assumptions that the χ^2 space is smooth and unimodal, that χ^2_ν is unity, that the spectral errors are all independent and Gaussian, and that the covariance between parameters is negligible (only the diagonal of the covariance matrix is used). It is worth noting that we have not rescaled the spectrum errors to artificially yield $\chi^2_\nu = 1$ in our calculation of the parameter errors we present.

With the exception of the σ_* uncertainties at low σ_* and S/N$_g$, the formal errors tend to underestimate the empirical errors; however, in the mean, the two measurements are most often consistent within a factor of two. We return to a discussion of these results from our repeat observations in the context of the parameter-recovery simulations presented in the next section.

7.5.2. Parameter-recovery Simulations

An industry-standard way to test the performance of fitting algorithms is to simulate data with known input parameters, apply the fitting algorithm to those data, and then compare the input and output parameters as a function of S/N (e.g., Bender 1990; Rix & White 1992; Westfall et al. 2011). Indeed,

[62] See also ppxf.capfit.cov_err in the pPXF Python package.

this is the invaluable first-order check of the validity of any fitting algorithm. These simulations have been performed multiple times with pPXF over the past >15 yr, including simulations specific to MaNGA spectra (Penny et al. 2016), as well as in the original pPXF papers (Cappellari & Emsellem 2004; Cappellari 2017).

In the most idealized approach to such simulations, the algorithm used to generate the model fit to the data is also used to generate the synthetic data, and the mock spectrum is constructed by one or more templates contained in the same set used for the fit. A key feature of these idealized tests is that any synthetic spectrum can be exactly reproduced by the provided template library, to within the limits of the random noise. Additionally, there is no resolution difference between the template and synthetic data and the stellar LOSVD is exactly parameterized (in the case of pPXF by a Gaussian or Gauss–Hermite function). In these idealized tests, pPXF has been shown to be robust, while the method, by design, penalizes the LOSVD toward a Gaussian when the data do not contain enough information to constrain the higher moments (Cappellari 2017).

This kind of idealized simulation is useful to assess possible intrinsic limitations of any method and it sets limits on what can be achieved, but it is not necessarily representative of what one should expect in real applications. In fact, in essentially any practical application of pPXF, the template library will be fundamentally limited in its ability to represent any given galaxy spectrum, the resolution estimation or resolution matching between the templates and galaxy data will have some uncertainty, and the stellar LOSVD will not be exactly represented by parametric forms. These model inaccuracies are very difficult to capture with simulated data in a meaningful way; i.e., in a way that is relevant to a specific application within the vast parameter space available.

Despite this criticism, parameter-recovery simulations, even if fully idealized, are still the best way to assess systematic error. As we noted in the previous section, repeat observations may provide the most direct assessment of random error but they are still limited by the fact that the true parameter values are unknown and that each observation could suffer from systematic error in an undetectable way. Therefore, the goal of our parameter-recovery simulations here is largely as a comparison to the statistics provided by the repeat observations and as a check for systematic errors.

In our simulations, we make a small step toward more realistic synthetic spectra by including the effect of template mismatch and poor resolution matching, in the sense that our mock solar spectrum is *not* included in the set of 42 MILES-HC templates we use for the pPXF fit. Moreover, the resolution of the templates is *not* perfectly matched to that of the mock spectrum. This implies that, even at infinite S/N, the pPXF fit will never perfectly reproduce the mock spectrum and some systematic deviations in the recovery should be expected.

To construct our synthetic MaNGA spectra, we convolve the BASS 2000 solar spectrum[63] with an input finely sampled Gaussian LOSVD, and then convolve again by a finely sampled Gaussian kernel to match the spectral resolution to a fiducial MaNGA resolution vector (Appendix A). The fiducial resolution vector is based on the median of the spectral resolution vectors for observation 7815-1902. We then integrate the synthetic spectrum over the pixels to match the spectral sampling of the MaNGA data and we add noise to meet a specific S/N level. We simulate noise in the spectrum by sampling from a Gaussian distribution that follows the mean trend of the flux variances with wavelength for observation 7815-1902. The noise trend is scaled to match the target S/N for each simulated spectrum. In a single simulation, one synthetic spectrum is constructed for each observed spectrum in a specific MaNGA RSS file (e.g., manga-7815-1902-LOGRSS.fits.gz) and each spectrum is fitted by the DAP using pPXF and the MILES-HC template library, as done for the real galaxies.

The results of our simulation are shown in Figure 20, using a panel layout that is identical to Figure 19 for the repeat observations. Also similar to our analysis of the repeat observations, we bin the data by S/N and σ_*; however, in this case we know and can bin by the *true* σ_* of the synthetic spectrum, where the number of synthetic spectra falling in each bin is given by $N_{\rm bin}$. Of course, instead of comparing paired results, these simulations compare the input and output kinematics.

Qualitatively, Figures 19 and 20 have many common features: χ^2_ν similarly increases toward high S/N, which reflects the inability of the MILES-HC library to exactly produce the solar spectrum; the random errors are similarly proportional to $(S/N)^{-1}$; and there is a similar transition of the random errors in σ_* to a constant value at low σ_* and S/N. In addition, the increasing inaccuracy of the formal velocity errors at $S/N < 10$ is qualitatively similar to the behavior of the repeat observations.

However, a trend of increasing systematic error emerges at low σ_* and S/N. This is due to the obvious fact that in this regime the σ_* uncertainties are so large that the best-fitting solution must often lie at the physical $\sigma_* = 0$ positivity boundary. In the limit where $\sigma_* = 0$ is the true value, the *average* σ_* is positively biased by an amount proportional to the σ_* errors, or inversely to the S/N. Small inaccuracies in the instrumental resolution of the template additionally contribute to the trends observed in this regime, given that even a small mismatch will produce a large effect on σ_* at very low $\sigma_{\rm obs}$. The modest bias in velocity for all σ_* is likely to be due to a slight error in the correction to 0 heliocentric velocity for one or both the BASS2000 spectrum and MILES-HC templates; this difference of $\lesssim$12 km s^{-1} is much smaller than the MaNGA pixel scale (70 km s^{-1}). Finally, we note that the formal errors do not exhibit the same trend toward underestimating the true error toward high S/N that is seen in the repeat observations.

For a more direct comparison with Figure 19, it is useful to rebin the data using the *output* σ_*, instead of the true value, as presented in Figure 21. Although it has relatively little effect on the velocity statistics, the choice of binned quantity substantially changes the σ_* statistics. This happens because now both the abscissa and ordinate are affected by the positivity boundary bias. In particular, any indication of systematic error in σ_* has vanished, reinforcing our claim that these simulations (i.e., Figure 20) provide key insights into systematic error, despite their other limitations compared to repeat observations. In addition, the transition to a constant fractional error in σ_* toward low S/N and the spread in $\delta\sigma_*/\langle\epsilon[\sigma_*]\rangle$ at low S/N are both more similar to the empirical behavior in Figure 19.

A distinct component of the repeat-observation statistics in Figure 19 that is not seen in our simulations is the increase in

[63] http://bass2000.obspm.fr/solar_spect.php

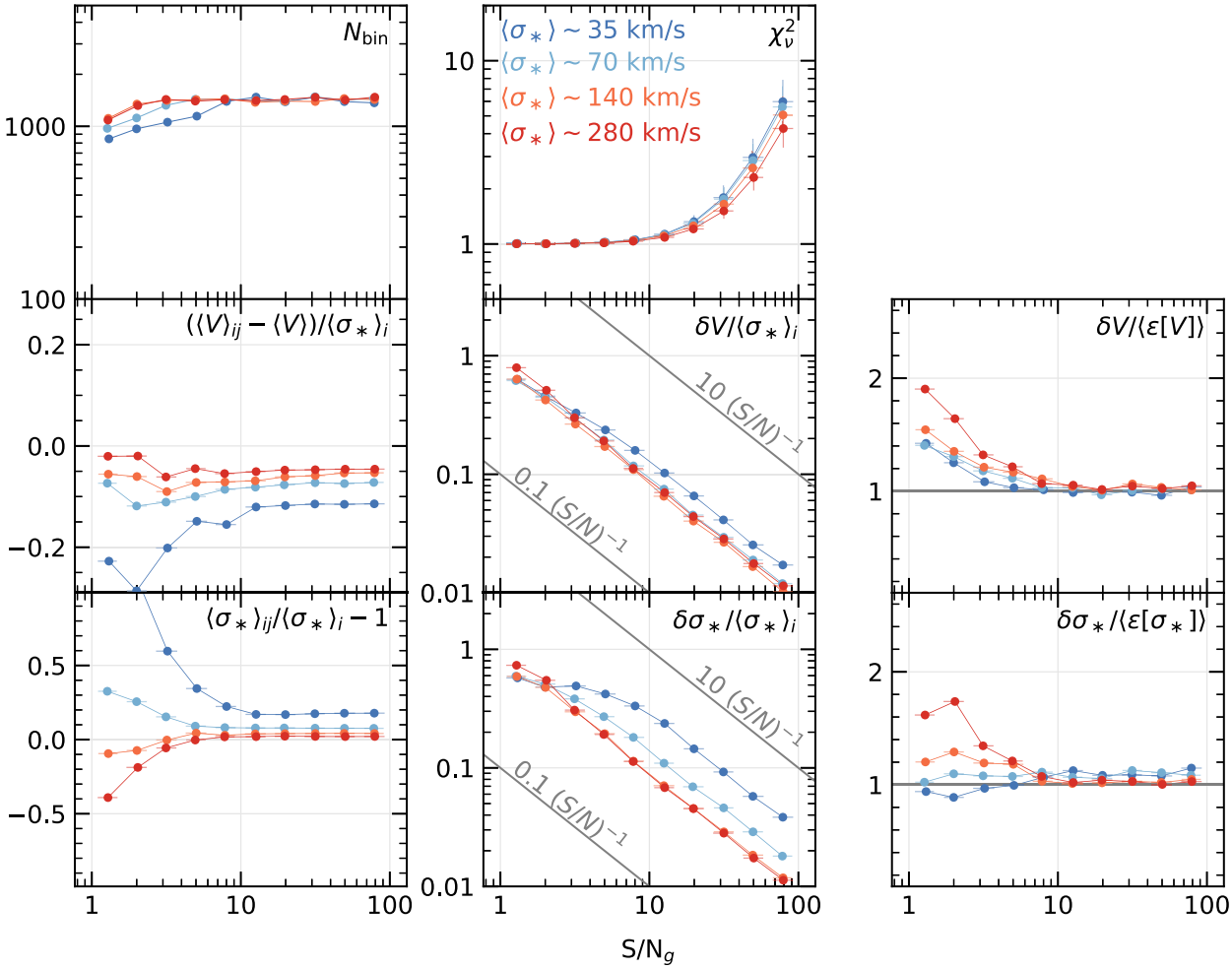

Figure 20. Nearly the same as Figure 19, but based on the idealized simulations discussed in Section 7.5.2. Instead of comparing the results from two fits, we compare the output kinematic measurement to the known input value used to construct each synthetic spectrum. The results are *binned by the known input value for* σ_* (see Figure 21).

the true error relative to the DAP-provided formal error toward high S/N. Despite the similar increase in χ^2_ν toward high S/N, the simulations show the formal errors are correct to within 10%–20% for S/N > 10. We expect this difference between the empirical and simulated data happens because of subtle issues in the repeatability of MaNGA's on-sky sampling, and propagation of the fiber data to the data cube leads to a more pronounced difference between the repeat data cubes at high S/N. In fact, in similar tests with the emission-line properties, as discussed in Section 3.3 of Belfiore et al. (2019), we find that these differences are fully consistent with the DRP-reported errors in the astrometric solution of MaNGA's dithered observations.

7.5.3. Algorithm Implications

The main algorithmic choice that we have made for DR15 based on the empirical and simulated performance of the DAP presented in this section is to bin the data to S/N ∼ 10 for our stellar kinematics measurements. This limits the number of σ_* values at the lowest $\sigma_* = 0$ boundary and the corresponding bias of the *average* value at lower S/N (Figure 20). We expect that this approach will also limit the difference between the true and DAP-reported errors in kinematics to ∼10%–20% at low S/N$_g$.

7.6. Flagging

Flagging of the DAP output stellar kinematics is somewhat minimal in DR15 with three basic flags applied (see Tables 8 and 9): (1) if a spectrum (binned or otherwise) does not meet either the binning or minimum S/N criterion for the fit (S/N > 1), the spaxel is masked and flagged as having NOVALUE; (2) in very rare cases (see Section 11), the fit returns a failure status flag indicating something went wrong in the fit, which we flag as FITFAILED; and (3) we use pPXF to impose boundaries restricting the viability of the derived kinematics, and set the NEARBOUND flag when the derived kinematics lie at one of these boundaries. For velocities, NEARBOUND is triggered if the velocity is within 1% of the imposed limits (i.e., $-2000 < v < -1980$ or $1980 < v < 2000$ km s^{-1}, with respect to the input redshift). This situation may indicate that the input redshift was not sufficiently accurate or there were problems with the input spectrum. For this reason, the fits and kinematics should not be trusted. For the velocity dispersion, NEARBOUND is triggered if the velocity dispersion is within 1% of the log of the range allowed values (i.e., $0.69 < \sigma < 0.74$ or $929.8 < \sigma < 1000$ km s^{-1}). This flag generally indicates that the galaxy velocity dispersion is lower than can be reliably measured. See further comments on the spaxel bit masks in Sections 11 and 12.

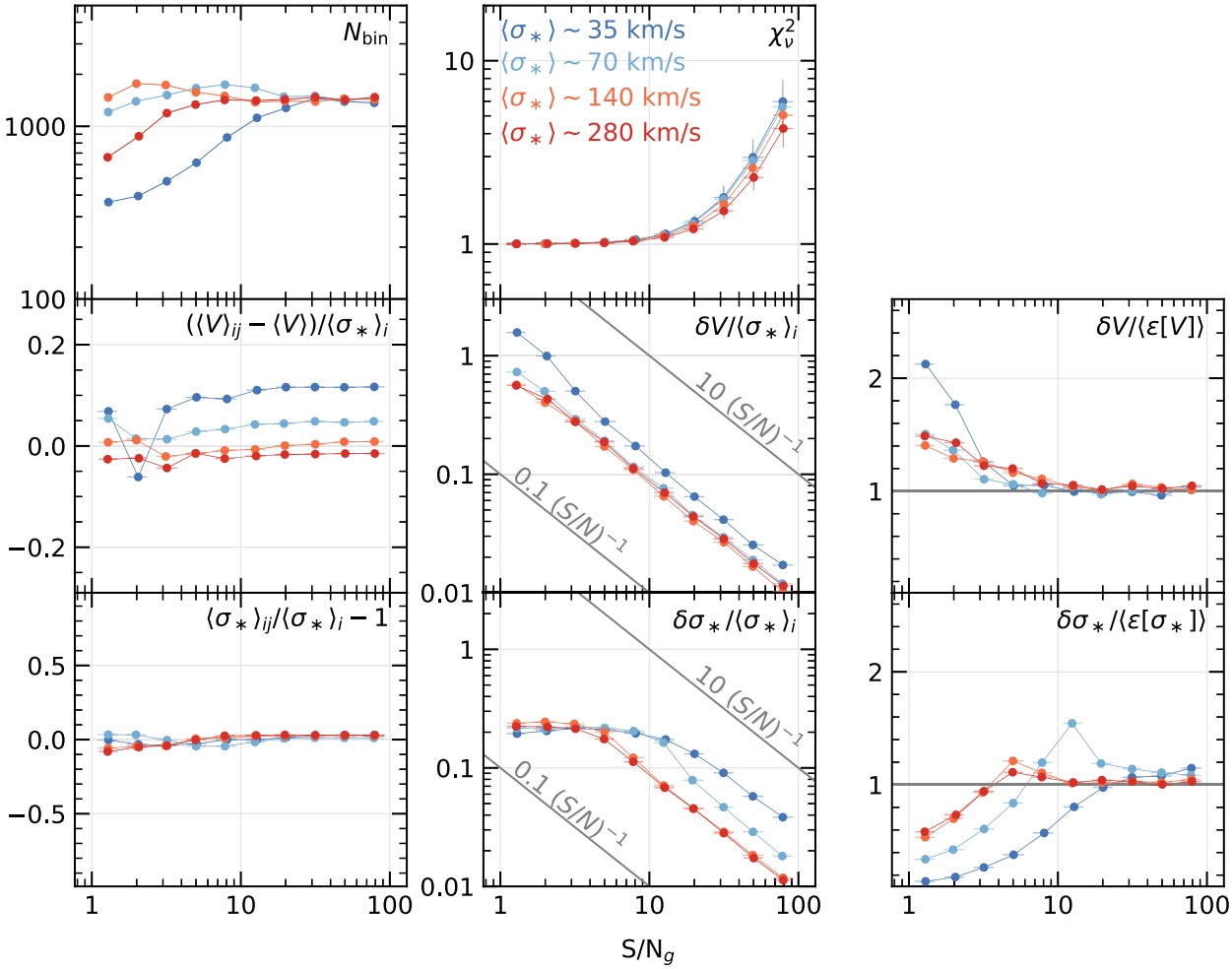

Figure 21. The same as Figure 20, except with the simulation results *binned by the measured output value for* σ_*.

7.7. Usage Guidance

Given the unusual format of our provided stellar velocity dispersion measurements, here we summarize and provide some usage guidance.

We emphasize again that the measurements of σ_* are *not* provided. Instead, we provide a "raw" measurement, $\sigma_{\rm obs}$—as determined by `pPXF` when fitting the MaNGA spectra with `MILES-HC` spectral templates (Section 5) including the intrinsic spectral-resolution offset—and a correction that removes the effects of the MaNGA-MILES resolution offset, $\delta\sigma_{\rm inst}$. The correction is applied by solving for σ_* in Equation (6). With a typical value of 33 km s^{-1} (Section 7.1.5), the correction is ∼10% and ∼5% when $\sigma_* = 70$ and 100 km s^{-1}, respectively. We have calculated spaxel-by-spaxel values of $\delta\sigma_{\rm inst}$ using the spectral resolution vectors provided for every MaNGA spaxel, and we recommend that they should be used instead of adopting a single correction for all spaxels. Convenience methods are provided in `Marvin` for this purpose.

Unfortunately, it is difficult to construct definitive criteria for when one should and should not trust any given determination of σ_* in DR15. We first emphasize that Figures 19 and 20 demonstrate that uncertainties in σ_* are a function of *both* σ_* and S/N$_g$, not just one or the other. Although we have binned our data to S/N$_g \gtrsim 10$ to minimize systematic error σ_*, not all bins will reach S/N$_g \sim 10$ (see Figure 9) and not all systematic errors are absent at S/N$_g \gtrsim 10$, particularly at low dispersion. Having said that, Figures 19 and 20 suggest that measurements of $\sigma_* \sim 35$ km s^{-1} are reasonable at S/N$_g > 20$; however, the occurrence of such measurements should be small, one can expect a systematic error of 10%–20%, and the random error will be large (30%–40%). A more appropriate limit, particularly given uncertainties in the LSF width, is likely higher such that we recommend extreme caution with measurements of $\sigma_* < 50$ km s^{-1} at S/N ∼ 10.

Based on Figures 20 and 21, we expect the random errors in our kinematics are valid given our binning to S/N$_g \gtrsim 10$. However, an important caveat to their accuracy is the increase in the standard deviation of the results from the repeat observations in the right-most panels of Figure 19. As we briefly mention and as discussed more fully by Belfiore et al. (2019), this increase in the standard deviation is consistent with the astrometric errors incurred when registering the individual, dithered MaNGA observations used to reconstruct each data cube. In other words, we expect this increase in error is caused by associating the kinematic measurement with a particular position on the galaxy, not an inaccuracy in the reported kinematic error as determined from the spectrum itself. We expect that users will want to associate each measurement with its spatial location and should therefore include this increase in the error, assuming they do not instead incorporate the astrometric errors directly. A rough, by-eye assessment of

Figure 19 suggests one should increase the error by a factor of $\log S/N_g$, valid over the range $10 \lesssim S/N_g \lesssim 100$.

Critically, none of the tests performed herein have considered the influence of errors in the LSF measurements—neither in MaNGA nor in the fitted template spectra—on the best-fit kinematics. For example, we have not yet presented a comparison of the MaNGA σ_* measurements with those made at higher spectral resolution (see Ryś et al. 2013; Falcón-Barroso et al. 2017). Although we continue to improve our measurements of the wavelength-dependent MaNGA LSF, we currently measure an error of ∼3%. This leads to ∼10% and ∼6% error in σ_* at ∼35 and ∼50 $\mathrm{km\,s^{-1}}$, respectively. Particularly at low σ_*, we encourage users to investigate how this error may or may not influence their analysis and subsequent inference.

Finally, we comment on the prevailing concern of how to treat measurements with $\sigma_{\rm obs} < \delta\sigma_{\rm inst}$; i.e., when σ_*^2 is negative. First note that of all of the measurements made for DR15, fewer than 2% show $\sigma_{\rm obs} < \delta\sigma_{\rm inst}$. In terms of spectra within a data cube, the median fraction of measurements with $\sigma_{\rm obs} < \delta\sigma_{\rm inst}$ is 0.3% and fewer than 5% of data cubes have $\sigma_{\rm obs} < \delta\sigma_{\rm inst}$ for more than 14% of their measurements. Even so, importantly, one should *not* consider such measurements equivalent to, nor a valid measurement of, $\sigma_* = 0$. We expect the predominant reason for measurements of $\sigma_{\rm obs} < \delta\sigma_{\rm inst}$ to be an effect of the S/N and spectral-resolution errors, as we discuss in detail in Section 7.4.3.

A minimalist approach is instead to simply ignore these measurements. As we discuss in Section 7.4.3, the measurements with $\sigma_{\rm obs} < \delta\sigma_{\rm inst}$ are very similar to the $\sigma_{\rm obs} = 0$ measurements made by the matched-resolution fit, with some exceptions at the lowest values of σ_*. Therefore, ignoring $\sigma_{\rm obs} < \delta\sigma_{\rm inst}$ measurements is nearly equivalent to ignoring measurements of $\sigma_{\rm obs}$ measurements from pPXF near the lower trust-region boundary on σ in its more common usage.

However, the minimalist approach ignores the possible benefits of including these $\sigma_{\rm obs} < \delta\sigma_{\rm inst}$ measurements, particularly for aggregation (e.g., the luminosity-weighted mean σ_* at fixed radius) and model fitting. In both cases, the results may be biased if one assumes a Gaussian error distribution when it is instead either truncated at $\sigma_* > 0$ or has a significant positive skew. Given this, we recommend including the negative σ_*^2 measurements when, e.g., constructing a mean by computing $(\sum_i \sigma^2_{{\rm obs},i} - \delta\sigma^2_{{\rm inst},i})^{1/2}$.[64] Such a computation can still lead to an imaginary number, which one can still ignore, but it provides a better accounting of the error distribution. In terms of model fitting, one can (for example) reformulate the merit (likelihood) function to incorporate both $\sigma_{\rm obs}$ and $\delta\sigma_{\rm inst}$—a reasonable estimate for the error in $\delta\sigma_{\rm inst}$ is 3%. The resolution difference between MaNGA and MILES is small (∼16%), such that the benefits of including $\sigma_{\rm obs} < \delta\sigma_{\rm inst}$ measurements may be limited but are still non-negligible.

8. Bandpass Integrals

Integrations over spectral regions are general to measurements of both emission-line moments (Section 9.7) and spectral indices (Section 10), which will be further discussed and defined in those sections. For clarity, we first introduce the formalism and nomenclature used when discussing these calculations below.

For a generic vector with values y_i sampled by $i = 0...N-1$ pixels at wavelengths λ_i, we calculate a bandpass integral via the discrete sum:

$$S(y) = \sum_i y_i\ dp_i\ d\lambda_i, \tag{8}$$

where dp_i is the fraction of the pixel, i, included in a given passband and $d\lambda_i$ is the wavelength step of the full pixel. For DR15, note that we set $dp_i = 0$ for any pixel within the band that is masked by the DRP, instead of attempting to replace the errant pixel by interpolation or sampling of a best-fit model spectrum; spectra with such masked pixels within the relevant passbands are flagged (Sections 9.7.1 and 10.4). Similarly, we calculate the mean of the spectrum as:

$$\langle y \rangle = \frac{S(y)}{S(1)}. \tag{9}$$

Finally, we calculate the weighted center of a passband as:

$$\langle \lambda \rangle = \frac{S(\lambda y)}{S(y)}. \tag{10}$$

The DAP currently only analyzes the spectra that are log-linearly sampled in wavelength meaning that $d\lambda_i \propto \lambda_i$.

When required for the measurement, we compute a linear pseudo-continuum using data in two passbands (sidebands) to either side (toward shorter—blue—and longer—red—wavelengths) of the main feature, as follows:

$$C_i = (\langle f \rangle_{\rm red} - \langle f \rangle_{\rm blue}) \frac{\lambda_i - \langle \lambda \rangle_{\rm blue}}{\langle \lambda \rangle_{\rm red} - \langle \lambda \rangle_{\rm blue}} + \langle f \rangle_{\rm blue}, \tag{11}$$

where f is the flux density and the "blue" and "red" subscripts denote calculations in the respective sidebands.

9. Emission-line Measurements

The DAP provides two sets of emission-line measurements. The results of Gaussian line-profile modeling are of primary interest, which will be the focus of this section. However, we also provide non-parametric measurements based on moment calculations of the continuum-subtracted data, as described in Section 9.7. Table 3 provides relevant data for the fitted emission lines with measurements in DR15; these data are referred to throughout this section.

The line-profile modeling performed by the DAP is the result of a second full-spectrum fit using pPXF that simultaneously models all emission-line features and re-optimizes the weights of the spectral templates to the underlying stellar continuum. Importantly, this second full-spectrum fit keeps the stellar kinematics fixed to the results from the first fit described in Section 7. For completeness, we provide a detailed description of this algorithm here. However, detailed assessments of our approach and suggested avenues for improvement are discussed in our companion paper, Belfiore et al. (2019).

In Belfiore et al. (2019), we pay particular attention to the effects of the continuum fit on the resulting emission-line properties and we assess the advantages of simultaneously fitting both components (gas and stars). This approach is similar to, and motivated by the success of, GANDALF (Sarzi et al. 2006, based on an early IDL version of pPXF); however, it is worth highlighting the following differences. Whereas

[64] With the addition of the luminosity weighting, this is the calculation we perform for the aggregated stellar velocity dispersion within $1R_e$ provided in the DAPall file; see Section 12.3.

Table 3
Emission-line Parameters

ID	Ion	λ_{rest}[a] (Å)	Ties[b]	Flux Ratio	Passbands (Å)		
					Main	Blue	Red
1	[O II]	3727.092	⋯	⋯	3716.3–3738.3[c]	3696.3–3716.3	3738.3–3758.3
2	[O II]	3729.875	k1	⋯	⋯	⋯	⋯
3	Hθ	3798.9826	⋯	⋯	3789.0–3809.0	3771.5–3791.5	3806.5–3826.5
4	Hη	3836.4790	⋯	⋯	3826.5–3846.5	3806.5–3826.5	3900.2–3920.2
5	[Ne III]	3869.86	⋯	⋯	3859.9–3879.9	3806.5–3826.5	3900.2–3920.2
6	Hζ	3890.1576	⋯	⋯	3880.2–3900.2	3806.5–3826.5	3900.2–3920.2
7	[Ne III]	3968.59	⋯	⋯	3958.6–3978.6	3938.6–3958.6	3978.6–3998.6
8	Hϵ	3971.2020	⋯	⋯	3961.2–3981.2	3941.2–3961.2	3981.2–4001.2
9	Hδ	4102.8991	⋯	⋯	4092.9–4112.9	4072.9–4092.9	4112.9–4132.9
10	Hγ	4341.691	⋯	⋯	4331.7–4351.7	4311.7–4331.7	4351.7–4371.7
11	[He II]	4687.015	⋯	⋯	4677.0–4697.0	4657.0–4677.0	4697.0–4717.0
12	Hβ	4862.691	⋯	⋯	4852.7–4872.7	4798.9–4838.9	4885.6–4925.6
13	[O III]	4960.295	a14	0.340	4950.3–4970.3	4930.3–4950.3	4970.3–4990.3
14	[O III]	5008.240	⋯	⋯	4998.2–5018.2	4978.2–4998.2	5018.2–5038.2
15	[He I]	5877.243	⋯	⋯	5867.2–5887.2	5847.2–5867.2	5887.2–5907.2
16	[O I]	6302.046	⋯	⋯	6292.0–6312.0	6272.0–6292.0	6312.0–6332.0
17	[O I]	6365.535	a16	0.328	6355.5–6375.5	6335.5–6355.5	6375.5–6395.5
18	[N II]	6549.86	a20	0.327	6542.9–6556.9	6483.0–6513.0	6623.0–6653.0
19	Hα	6564.632	⋯	⋯	6557.6–6571.6	6483.0–6513.0	6623.0–6653.0
20	[N II]	6585.271	⋯	⋯	6575.3–6595.3	6483.0–6513.0	6623.0–6653.0
21	[S II]	6718.294	⋯	⋯	6711.3–6725.3	6673.0–6703.0	6748.0–6778.0
22	[S II]	6732.674	⋯	⋯	6725.7–6739.7	6673.0–6703.0	6748.0–6778.0

Notes.
[a] Ritz wavelengths in vacuum from the National Institute of Standards and Technology (NIST; http://physics.nist.gov/PhysRefData/ASD/Html/help.html).
[b] The velocities of *all* lines are tied to one another; "k*n*" signifies the line has all its kinematics (V, σ) tied to line with ID n; "a*n*" signifies that the line has all its parameters tied to line n with a fixed flux ratio.
[c] The [O II] doublet is unresolved. Our moment measurements adopt a primary band that brackets both emission lines. One must sum the Gaussian results for both lines in the doublet when comparing to the single, zeroth-moment measurement.

pPXF treats both the gas and stellar spectra in exactly the same way (as spectral templates; see Cappellari 2017, Section 3.6), GANDALF models the emission lines as Gaussian functions that are added to the stellar spectra within the code. The pPXF approach for the gas emission was made possible by the analytic Fourier convolution (Cappellari 2017) introduced in version 6.0 of pPXF package in Python. This approach is more general and allows for greater flexibility to the user, but it requires some initial setup. Therefore, we go into some detail about the construction of the emission-line templates in Section 9.1.

For DR15, the template library used for the stellar component (MILES-HC) is also used during the stellar-kinematics fit. However, even though we expect the MILES-HC library to provide the best measurements of the stellar kinematics, they prove more limiting in the emission-line modeling. In particular, the spectral range of the MILES library limits the emission lines that can be modeled (Figure 2), prohibiting models of the [S III] lines at 9071.1 and 9533.2 Å. In addition, the unrestricted optimization of the MILES-HC templates ignores all of the physics that we understand about stellar mixes relevant to how stars are formed and evolve in galaxies. Belfiore et al. (2019) explore the use of stellar-population-synthesis templates with larger spectral coverage in a development version of the DAP that allows for different templates to be used for the stellar kinematics and the emission-line modeling.

The spatial binning of the data (Section 6) for DR15 is primarily driven by the desire to meet a S/N threshold that minimizes any systematic errors in the stellar kinematics (Section 7.5). It is reasonable then that the definition of S/N used for binning the data is based on the broad-band S/N within the SDSS g filter. However, this means that the spatial binning applied in Section 6 is likely to be inappropriate for emission-line focused science. It is indeed common-place nowadays for emission lines to be fit with a different binning scheme than used for analyses of the broad-band continuum. We have implemented a similar strategy in the DAP, where the stellar kinematics are performed on the Voronoi-binned data and the emission-line modeling can be performed on individual spaxels. We refer to this as the "hybrid" binning output. However, it is also useful to have the emission lines fit using the same spectra as used for the stellar continuum, particularly if the goal is simply to subtract them for subsequent analysis (e.g., Wilkinson et al. 2017). This difference between fitting the emission lines in the binned spectra or the individual spaxel data is the primary distinction between the two types of data provided by the DAP for DR15 (DAPTYPE is VOR10-GAU-MILESHC or HYB10-GAU-MILESHC; see Section 12). Sections 9.2 and 9.3 separately describe the fit iterations in these two methods for the sake of clarity.

9.1. Emission-line Template Construction

Similar to GANDALF, the DAP sets how the emission-line templates are constructed via an input file. We are free to define (1) the functional form used to construct the emission line, (2) any lines that should have a fixed flux ratio, and (3) how line

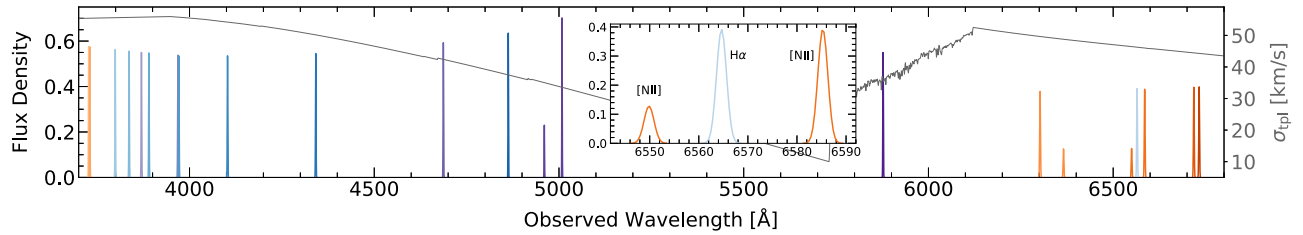

Figure 22. Emission-line templates constructed for fitting observation 7815-3702, as described in Section 9.1, for the emission lines listed in Table 3. Spectra associated with each template are colored according to the line groups defined by Belfiore et al. (2019): Balmer lines in blue, low-ionization lines ([O II], [O I], [N II], and [S II]) in orange, and high-ionization lines ([Ne III], He II, [O III], and He I) in purple (see Figure 2). The inset shows the [N II]+Hα complex. Note the [N II] doublet features are components of the same template spectrum such that the relative kinematics and flux ratio of the two lines are held fixed. The resolution of the template features, $\sigma_{\rm tpl}$, is plotted in gray following the right ordinate.

parameters should be tied together to force all lines to have the same velocity and/or velocity dispersion. Table 3 provides the fixed flux ratios adopted for the [O III], [O I], and [N II] lines in DR15, as well as how lines are tied kinematically. We force the velocities of all lines to be identical, and we force the velocity dispersion of the [O II], [O III], [O I], and [N II] doublet features to be the same. Belfiore et al. (2019) explores the effects of different tying strategies on the emission-line properties, such as grouping lines by ionization potential (Figure 2).

The emission-line templates are constructed once per fitted data cube. Each emission feature is included in the relevant template with a known flux (typically unity), line center (at the rest wavelength), and width. The width of the line is set based on the minimum instrumental dispersion over all spatial positions in the data cube, as given by the DRP-provided spectral-resolution measurements. However, we apply a quadrature offset, such that the minimum line-profile dispersion is 10 km s^{-1}. That is, we first compute the minimum instrumental dispersion for each wavelength, λ, over all spatial positions, (x, y):

$$\sigma_{\rm inst,min}(\lambda) = \min_{xy}\left[\frac{c}{\sqrt{8\ln 2}}\,R_{xy}^{-1}(\lambda)\right], \quad (12)$$

where the resolution at a given spatial position, $R_{xy}(\lambda)$, is sampled at the observed wavelengths, $\lambda = \lambda_{\rm tpl}(1+z_0)$ (z_0 is the cosmological redshift; see Section 7.1.4). We then apply the quadrature offset:

$$\sigma^2_{\rm tpl}(\lambda) = \sigma^2_{\rm inst,min}(\lambda) - \min_\lambda[\sigma^2_{\rm inst,min}] + 10^2. \quad (13)$$

By significantly lowering the width of the lines in the template, we ensure that the pPXF fit yields a non-zero dispersion measurement for *all* lines, accommodating uncertainties in and variation of the LSF within the data cube. For example, the emission-line templates and the relevant $\sigma_{\rm tpl}$ used during the fit of observation 7815-3702 are shown in Figure 22.

Lines that are forced to have the same flux ratio can be modeled by including them in a single template. For example, see the [N II] lines in the inset panel of Figure 22. The fixed flux ratio is then achieved naturally in the linear optimization of the template weights performed by pPXF. All other lines must have their own unique template, such that there are 19 emission-line templates used in DR15. The DAP assigns each emission-line template to a velocity and velocity-dispersion group, which is used to construct the initial parameter guesses and the parameter tying structure used by pPXF. Finally, the emission-line templates are combined with the stellar templates and assigned kinematic components, one component for the stars and separate components as needed for the kinematic parameters to be fit for the emission lines.

9.1.1. Rigorous Sub-pixel Emission-line Templates

As stated previously, the quadrature offset applied to the emission-line profiles in our emission-line templates is to ensure that we measure a non-zero velocity dispersion for *all* emission lines when fitting the MaNGA galaxy spectra. However, in making the template lines narrower, we risk under-sampling the emission features. Indeed, if the lines are too narrow, then the use of a Gaussian function directly to define the line profiles would lead to very inaccurate results. In general, therefore, we account for this issue in a mathematically rigorous manner by defining the template emission-line profile directly in Fourier space (Cappellari 2017):[65]

$$\mathcal{L}(x) = \mathcal{F}^{-1}\left[\exp\left(-\frac{\omega^2\sigma_x^2}{2} - i\,\omega\,x_0\right)\right], \quad (14)$$

where the expression within the square brackets is the analytic Fourier transform of a Gaussian with dispersion σ_x and center x_0 in the same logarithmically spaced spectral pixels as for the stellar templates. Here, $\mathcal{F}^{-1}$ is the inverse of the *discrete* Fourier transform of real input. Importantly, the calculation of the inverse Fourier transform is required only once per emission line, meaning that the overhead compared to sampling a Gaussian function is entirely negligible.

With this definition, the convolution of the emission-line template within pPXF produces, by construction, a numerically accurate Gaussian regardless of the width of the emission line. Indeed, this approach even allows one to use mathematically correct Dirac delta functions as emission-line profiles in pPXF by simply setting $\sigma_x = 0$ in Equation (14). When the emission line is well sampled, namely when $\sigma_x \gtrsim 1$, then Equation (14) reduces to a normalized Gaussian function that *sums* to unity.

In general, for maximum accuracy, one will want to include the integration of the emission lines over the spectral pixels before fitting the observed spectrum. Ignoring this effect is equivalent to adopting an instrumental dispersion ≈3% lower than the "true" value one would measure with a well-sampled LSF (Figure 4 of Cappellari 2017). In this case, the emission-

[65] One can compute this profile using the Python function ppxf.ppxf_util.Gaussian provided with the pPXF software package. For the results from Equations (14) or (15), set pixel=False or pixel=True, respectively.

line profiles becomes

$$\mathcal{L}(x) = \mathcal{F}^{-1}\left[\exp\left(-\frac{\omega^2\sigma_x^2}{2} - i\,\omega\,x_0\right)\mathrm{sinc}\left(\frac{\omega}{2\pi}\right)\right], \quad (15)$$

where the expression in square brackets now represents the analytic Fourier transform of a Gaussian convolved with an unitary pixel, with the sinc function being the Fourier transform of a unitary box function. Like before, this line profile produces a mathematically accurate Gaussian integrated within the pixels when convolved with the ionized-gas LOSVD in the pPXF fit.

9.1.2. Usage Notes

We make the following two notes regarding the practical use of this method of constructing the emission-line profiles in the DAP.

First, at least for the data in DR15, the emission-line templates are constructed on the same wavelength grid as the stellar templates, meaning that the pixel sampling is a factor of four more frequent than the MaNGA galaxy spectra (Section 7.1.1; i.e., the pixel width is 17.3 km s^{-1}). Even so, the definition of the Gaussian profiles for the emission-lines directly in Fourier space will be critical to robust measurements for lines falling in spectral regions at our minimum $\sigma_{\rm tpl}$ of 10 km s^{-1}.

Second, although not relevant to the example templates shown in Figure 22 given the form of $\sigma_{\rm tpl}$, the line profiles generated using Equation (15) will show characteristic Fourier ringing when $\sigma_{\rm tpl}$ is much smaller than the size of the pixel. Although germane to the goal of a mathematically rigorous convolution, these features may remain in the final fit if the best-fit, convolved profile is significantly under-sampled. We know of no specific examples of this in the DAP results provided in DR15; however, we raise the issue here and continue to assess its influence, if any, in general use of the DAP.

9.2. Fit Iterations: Remapping from Binned Spectra to Individual Spaxels, the Hybrid Scheme

Our "hybrid" approach to emission-line fitting refers to the combination of having the stellar kinematics determined using the binned spectra, whereas the combined stellar-continuum +emission-line fit is performed for individual spaxels. In this approach, the emission-line fitter runs through three fit iterations:

Step (1): We fit the binned spectra assuming that the ionized gas is part of a single kinematic component; i.e., all lines have the same velocity and velocity dispersion. Only multiplicative polynomials are used to adjust the stellar-continuum model such that the absorption-line equivalent-width distribution from the stellar library is preserved. The multiplicative polynomial is only applied to the stellar templates, not the emission-line templates, to maintain any flux ratio between line doublets. The spectra are fit twice, once on the input binned spectra and then again after removing 3σ outliers. When fitting each binned spectrum, the weights of *all* templates (MILES-HC for DR15) are optimized; i.e., the emission-line module does not implement the NZT approach used by the stellar-kinematics module (Sections 7.1.3 and 7.4.1). Instead, the results of this fit are used to set stellar-continuum template weights and initial guess kinematics that are *unique to each binned spectrum* and applied to the spaxels associated with that bin. Spaxels are associated to the binned spectra by simple on-sky proximity.[66] The stellar continuum used in the fit to the individual spaxels is determine by re-weighting the optimal template determined for the associated binned spectrum. That is, in all remaining iterations, the templates used by pPXF to model the spectra consist of a single, optimized template for the stellar continuum and the full set of emission-line templates.

Step (2): We fit the individual spaxels, again assuming the ionized gas is in a single kinematic component. The fit includes a 3σ rejection iteration and the initial starting guess for the next iteration are updated from the previous fit.

Step (3): Finally, the individual spaxels are fitted again, however, this iteration assigns the kinematic components as dictated by the input file; see Section 9.1 and Table 3. This iteration *does not* include a rejection iteration.

9.3. Fit Iterations: Binned Spectra, No Remapping

When the emission-line fits are provided for the binned spectra and are not remapped to the individual spaxels, the emission-line fitter runs through two fit iterations. These iterations are virtually identical to the last two iterations in Section 9.2 with the additional step of obtaining the optimal stellar-continuum template in the first iteration. In the interest of completeness, the two iteration steps are:

Step (1): We fit the binned spectra assuming the ionized gas is part of a single kinematic component and including a 3σ rejection iteration. This iteration determines the optimal stellar-continuum template and initial starting guesses for the second iteration (see Step 1 in Section 9.2).

Step (2): The binned spectra are fitted again, however, this iteration assigns the kinematic components as dictated by the input file; see Section 9.1 and Table 3. This iteration *does not* include a rejection iteration (see Step 3 in Section 9.2).

9.4. Emission-line Modeling Results

Unlike the stellar-continuum fit, there is a non-trivial step in parsing the results of the pPXF fit into the relevant quantities for the emission-line modeling. The nuances of this parsing are discussed here, and particularly important usage implications are further discussed in Section 12.

Our general approach is to provide flux, equivalent-width, velocity (see Section 7.1.4), and velocity-dispersion measurements for *all* lines listed in Table 3, even for quantities that are constrained by multiple lines. For fluxes, our convention is to construct the emission-line templates such that the *integral* of the line is unity, except for those lines with a fixed flux ratio (Section 9.1). However, the convolution performed by pPXF conserves the *sum* of the template pixel values, not its integral. Therefore, to calculation the best-fitting flux for each emission line based on the input flux of line l, $F_{{\rm tpl},l}$, and the

[66] This approach is a development hold-over. Of course, we know exactly which spaxels were combined into each bin, meaning that this step is unnecessary and can actually lead to incorrectly matching a spaxel to the relevant bin. In recent updates to the DAP (MPL-8; version 2.3.0), we associate each spaxel directly to the correct parent bin.

optimized weight of the associated template j, w_j, we account for the effect of the convolution convention on the line *integral* following:

$$F_l = w_j F_{\mathrm{tpl},l}(1 + z_j); \tag{16}$$

note that $F_{\mathrm{tpl},l} = 1$ except where given otherwise in Table 3. We emphasize that the need for Equation (16) is an artifact of the template construction and the pPXF convolution.

Critically, note that all of the emission-line fluxes have been corrected for Galactic foreground extinction because the spectra are appropriately de-reddened (Section 6.5) before being modeled. However, we have not attempted to correct any of the line fluxes for attenuation by the dust content of the target galaxy itself.

Emission-line equivalent widths based on the line-profile modeling results are calculated as:

$$\mathrm{EW}_l = \frac{1}{1 + z_l} \frac{F_l}{C_l}, \tag{17}$$

where F_l is from Equation (16) and C_l is the pseudo-continuum of the observed spectrum (see Equation (11)) interpolated at the line center, $\lambda = (1 + z_l)\lambda_{\mathrm{rest}}$, using the definition of the sidebands in Table 3 (see the discussion of the passbands used here and in the non-parametric calculations in Section 9.7). Note that, in a slight modification to Equation (11), the computation of C_l in Equation (17) does not use the passband-integrated mean flux in each sideband (e.g., $\langle f \rangle_{\mathrm{red}}$), but instead uses a simple median of the flux within the passband. Also note that although Equation (17) is defined such that emission and absorption yield, respectively, positive and negative EW, the non-negative constraint on the template weights imposed by pPXF means that the model-based equivalent widths *cannot* be negative. Finally, note that the factor of $(1 + z_l)$ in Equation (17) is needed to convert our *observed-frame* equivalent-width measurement to the *rest frame*.

The velocity and velocity dispersion of each line is determined by its associated kinematic component. The velocity measurements are offset by the cosmological redshift of the galaxy in a manner identical to the method used for the stellar velocities described in Section 7.1.4. Also similar to the stellar kinematics, the gas velocity dispersions are provided as one would measure directly from the spectrum and must be corrected by the user using the provided instrumental dispersion (see Section 9.5).

Errors are also provided for each line flux, equivalent-width, velocity, and velocity dispersion regardless of whether or not the parameter was fit independently. This has important implications for the propagation of errors in averaged quantities. For example, *all* of the velocity errors for the lines in a given spectrum are the same because the velocities of all lines are tied during the fit, meaning that one cannot combine the velocities of multiple lines in a nominal way to improve the velocity precision.

9.5. Velocity Dispersion Corrections

For the ionized-gas velocity dispersions, the instrumental resolution $\sigma_{\mathrm{inst},i}$ at the wavelength of each line i is provided (Section 12; Appendix C) such that the corrected velocity dispersion of each line is

$$\sigma^2_{l,i} = \sigma^2_{\mathrm{obs},i} - \sigma^2_{\mathrm{inst},i}, \tag{18}$$

where σ^2_{obs} is the velocity dispersion of the line in the observed spectrum; the DAP MAPS files provide σ^2_{obs} and σ^2_{inst}.

9.6. Flagging of Modeling Results

Flagging of the emission-line model fitting results (see Tables 8 and 9) are similar to those used for the stellar kinematics, including the caveat that they are currently rather limited. The NOVALUE and NEARBOUND flags have the same meaning; the criteria used to flag data as NEARBOUND are the same as used for the stellar kinematics. In addition to a core failure of the fitting algorithm, the FITFAILED flag is also used to signify an error in the computation of the formal errors in the best-fit parameters.

9.7. Non-parametric Emission-line Measurements

Although more precise measurements of the emission-line properties are determined by our Gaussian modeling, it is useful to perform direct, non-parametric measurements of the emission features. In particular, these measurements provide initial estimates for the model optimization, and they lead to valuable assessments of the model results that can identify catastrophic errors and non-Gaussianity of the line profiles. Therefore, we compute zeroth-, first-, and second-order moments of the emission-line profiles using continuum-subtracted spectra, and we combine the zeroth moment and measurements of the local continuum to provide non-parametric emission-line equivalent widths.

In the calculation of the emission-line moments, the model continuum is subtracted from the data to produce emission-line-only spectra. Even with this subtraction, however, coherent deviations of the baseline local to each emission line may remain. Therefore, we also subtract a linear baseline below each emission feature following from Equation (11) and two sidebands to either side of the emission feature (see below). The definitions of the main passband over which the moments are calculated, as well as the two sideband definitions, are provided in Table 3. These same passbands are used in the equivalent-width calculation (see Section 9.4).

The definition of the emission-line passbands are taken from Yan et al. (2006, Table 3) for the majority of the strong lines. For the additional lines in Table 3, we adopt 20 Å passbands and place the sidebands directly to either side of the main passband. The exceptions to this are when the sidebands overlap with the main passband of other lines we intend to measure. In those cases, the sideband limits are adjusted as necessary to regions that should be free from emission. Other specific changes relative to the definitions from Yan et al. (2006) are that we apply a slight shift to the main passband for Hβ to center the passband on the line and we use a narrower Hα main passband to avoid overlap with the [N II] lines. Given their small separation, a number of line groups use the same sidebands; these groups are (1) Hη, Hζ, and [Ne III]; (2) Hα and [N II]; and (3) the two [S II] lines. Also note that only one main passband is used for the [O II] doublet because these lines are unresolved by the BOSS spectrographs.[67] All passbands

[67] Space is actually allocated for two non-parametric measurements of the [O II] doublet in the MAPS file; however, this is simply to establish a symmetry between the data format of the non-parametric and Gaussian-fit results. More detail is provided in Section 12 and Appendix C.

are defined at rest wavelengths and the passbands are appropriately redshifted such that all measurements are performed on the observed spectra.

Following the definition in Equation (8), we calculate the zeroth moment of the each line in flux units as:

$$\mu_{l,0} = S(f_l - B), \tag{19}$$

where f_l is the continuum-subtracted, emission-line-only spectrum and B is the linear baseline. The linear baseline is determined using Equation (11), except that we use the unweighted center of each of the "blue" and "red" passbands, instead of the spectrum-weighted center, because the passband-integrated flux of f_l is nearly zero. We then calculate the first and second moments in units of km s^{-1} as:

$$\mu_{l,1} = \frac{c}{\mu_{l,0}} S\left[\left(\frac{\lambda}{\lambda_{\rm rest}} - 1\right)(f_l - B)\right], \tag{20}$$

and

$$\mu_{l,1}^2 + \mu_{l,2}^2 = \frac{c^2}{\mu_{l,0}} S\left[\left(\frac{\lambda}{\lambda_{\rm rest}} - 1\right)^2 (f_l - B)\right], \tag{21}$$

respectively; $\mu_{l,1}$ is equivalent to a non-parametric Doppler shift of the line and $\mu_{l,2}$ is the dispersion of the line profile about that Doppler shift. The calculation of the non-parametric equivalent width is identical to that used by the line-profile modeling, except that we set $F_l = \mu_{l,0}$ in Equation (17). Although we have provided errors for the non-parametric measurements, they have not been as well vetted as the model-based results (see Belfiore et al. 2019). We expect their accuracy to be similar to what we find for the spectral indices (see Sections 10.3.1 and 10.3.2). However, we have not calibrated their accuracy, as done for the spectral indices, such that the values provided in DR15 should be treated with caution or simply ignored.

The non-parametric emission-line measurements are performed twice (Figure 3). The measurements are first performed before the emission-line modeling using z_0 as the redshift *for all spectra in the data cube* and the best-fitting model from the stellar-kinematics module as the continuum. The first moment of the Hα line from these measurements is used as the initial guess for the velocity in the emission-line modeling of each spectrum. The measurements are performed a second time after the emission-line modeling and they serve two purposes: (1) the emission-line modeling includes a simultaneous adjustment of the stellar continuum, such that the continuum from the first full-spectrum fit (Section 7) can be different from the continuum used by the emission-line modeling—redoing the measurements after the emission-line modeling is finished allows us to force the continuum used in both modules to be identical;[68] and (2) instead of a single redshift for all spectra, we use the best-fitting emission-line velocities (tied for all lines in a given spectrum in DR15) as the Doppler shift, again allowing us to minimize the systematic differences in the non-parametric and Gaussian-fit measurements. Only the second set of measurements are provided in the main DAP output files (Section 12; Appendix C).

9.7.1. Flagging

Through these non-parametric calculations, we set a number of flags (see Table 8). (1) If any of the passbands are empty, then no measurement is made and the NOVALUE maskbit is set. (2) When constructing the emission-line-only spectrum, discontinuities in the continuum will occur in the transition between regions that are and are not fit by the relevant full-spectrum-fitting module. If such a discontinuity lands within or between passbands, then no measurement is made and the FITFAILED maskbit is set. This is not really a concern for DR15, but is listed here for completeness. (3) If the measurement is made in a region without the continuum subtracted, then the measurement is provided but the NOCORRECTION maskbit is set. (4) If there are masked pixels within any of the passbands, then the measurement is provided but the UNRELIABLE maskbit is set. (5) In rare cases, the calculation of the moments or equivalent widths requires a division by zero; these cases are flagged as MATHERROR.

10. Spectral Indices

We use the term spectral index generally to refer to a measurement of a specific continuum feature in a galaxy spectrum made manifest by its stellar population. Because the spectral indices are meant to quantify features on the continua of galaxy spectra, all of our spectral-index measurements are made after first subtracting the best-fit emission-line model from the observed spectrum. The spectral indices provided by the DAP fall into the following two groups.

Absorption-line indices—such as those defined in the Lick Index system (Trager et al. 1998)—are measured similarly to emission-line equivalent widths. These indices are defined to measure the strength of absorption features predominantly associated with one, or a small number, of atoms or molecules present in stellar atmospheres. Following Equations (2) and (3) from Worthey (1994) and our Equation (8), we compute absorption-line indices as:

$$\mathcal{I}_a = \begin{cases} \dfrac{1}{1+z} S(1 - f_c/C), & \text{for Å units} \\ -2.5 \log [S(f_c/C)], & \text{for magnitude units,} \end{cases} \tag{22}$$

where f_c is the continuum-only spectrum determined by subtracting the best-fitting emission-line model and C is calculated using Equation (11) and the continuum-only spectrum. As indicated by Equation (22), absorption-line indices can be measured in magnitude units or Å and are constructed such that features seen in absorption relative to the pseudo-continuum yield positive index values. Except for some subtle difference in the definition, EW$_l \approx -\mathcal{I}_a$ when $\mathcal{I}_a$ has Å units and both use the same passbands. The first 43 rows in Table 4 provide the three passbands used in measuring $\mathcal{I}_a$ by the DAP in DR15; the table includes the units of each index and whether the passbands were defined for air or vacuum wavelengths.

Bandhead, or color, indices are measurements that simply compare the flux in two passbands, usually placed to either side of significant continuum-break features due to the atomic or

[68] The baseline, B, is always included, meaning that the *local* continuum can still be different between the non-parametric and Gaussian-fit results.

Table 4
Spectral-index Parameters

ID	Index	Passbands (Å)			Medium	Units	References	ε	$\delta\epsilon$
		Main	Blue	Red					
1	CN1	4142.125–4177.125	4080.125–4117.625	4244.125–4284.125	Air	mag	(1)	−0.54	1.3
2	CN2	4142.125–4177.125	4083.875–4096.375	4244.125–4284.125	Air	mag	(1)	−0.45	1.6
3	Ca4227	4222.250–4234.750	4211.000–4219.750	4241.000–4251.000	Air	Å	(1)	0.79	1.4
4	G4300	4281.375–4316.375	4266.375–4282.625	4318.875–4335.125	Air	Å	(1)	1.01	1.4
5	Fe4383	4369.125–4420.375	4359.125–4370.375	4442.875–4455.375	Air	Å	(1)	1.13	1.9
6	Ca4455	4452.125–4474.625	4445.875–4454.625	4477.125–4492.125	Air	Å	(1)	0.83	1.4
7	Fe4531	4514.250–4559.250	4504.250–4514.250	4560.500–4579.250	Air	Å	(1)	1.03	1.7
8	C24668	4634.000–4720.250	4611.500–4630.250	4742.750–4756.500	Air	Å	(1)	1.18	1.9
9	Hβ	4847.875–4876.625	4827.875–4847.875	4876.625–4891.625	Air	Å	(1)	0.76	1.2
10	Fe5015	4977.750–5054.000	4946.500–4977.750	5054.000–5065.250	Air	Å	(1)	1.13	1.9
11	Mg1	5069.125–5134.125	4895.125–4957.625	5301.125–5366.125	Air	mag	(1)	−0.82	1.4
12	Mg2	5154.125–5196.625	4895.125–4957.625	5301.125–5366.125	Air	mag	(1)	−0.73	1.3
13	Mgb	5160.125–5192.625	5142.625–5161.375	5191.375–5206.375	Air	Å	(1)	0.81	1.4
14	Fe5270	5245.650–5285.650	5233.150–5248.150	5285.650–5318.150	Air	Å	(1)	0.85	1.4
15	Fe5335	5312.125–5352.125	5304.625–5315.875	5353.375–5363.375	Air	Å	(1)	0.96	1.7
16	Fe5406	5387.500–5415.000	5376.250–5387.500	5415.000–5425.000	Air	Å	(1)	0.87	1.7
17	Fe5709	5696.625–5720.375	5672.875–5696.625	5722.875–5736.625	Air	Å	(1)	0.87	1.6
18	Fe5782	5776.625–5796.625	5765.375–5775.375	5797.875–5811.625	Air	Å	(1)	0.84	1.5
19	NaD	5876.875–5909.375	5860.625–5875.625	5922.125–5948.125	Air	Å	(1)	0.90	1.5
20	TiO1	5936.625–5994.125	5816.625–5849.125	6038.625–6103.625	Air	mag	(1)	−0.69	1.7
21	TiO2	6189.625–6272.125	6066.625–6141.625	6372.625–6415.125	Air	mag	(1)	−0.80	1.8
22	Hδ_A	4083.500–4122.250	4041.600–4079.750	4128.500–4161.000	Air	Å	(2)	1.02	1.2
23	Hγ_A	4319.750–4363.500	4283.500–4319.750	4367.250–4419.750	Air	Å	(2)	1.00	1.3
24	Hδ_F	4091.000–4112.250	4057.250–4088.500	4114.750–4137.250	Air	Å	(2)	0.83	1.1
25	Hγ_F	4331.250–4352.250	4283.500–4319.750	4354.750–4384.750	Air	Å	(2)	0.75	1.1
26	CaHK	3899.5–4003.5	3806.5–3833.8	4020.7–4052.4	Air	Å	(3)	1.43	1.5
27	Ca II1[a,b]	8484.0–8513.0	8474.0–8484.0	8563.0–8577.0	Air	Å	(4)	…	…
28	Ca II2[a,b]	8522.0–8562.0	8474.0–8484.0	8563.0–8577.0	Air	Å	(4)	…	…
29	Ca II3[a,b]	8642.0–8682.0	8619.0–8642.0	8700.0–8725.0	Air	Å	(4)	…	…
30	Pa17[a,b]	8461.0–8474.0	8474.0–8484.0	8563.0–8577.0	Air	Å	(4)	…	…
31	Pa14[a,b]	8577.0–8619.0	8563.0–8577.0	8619.0–8642.0	Air	Å	(4)	…	…
32	Pa12[a,b]	8730.0–8772.0	8700.0–8725.0	8776.0–8792.0	Air	Å	(4)	…	…
33	MgICvD	5165.0–5220.0	5125.0–5165.0	5220.0–5260.0	Vacuum	Å	(5)	0.94	1.4
34	NaICvD[b]	8177.0–8205.0	8170.0–8177.0	8205.0–8215.0	Vacuum	Å	(5)	…	…
35	MgIIR[b]	8801.9–8816.9	8777.4–8789.4	8847.4–8857.4	Vacuum	Å	(5)	…	…
36	FeHCvD[b]	9905.0–9935.0	9855.0–9880.0	9940.0–9970.0	Vacuum	Å	(5)	…	…
37	NaI[b]	8168.500–8234.125	8150.000–8168.400	8235.250–8250.000	Air	Å	(6)	…	…
38	bTiO	4758.500–4800.000	4742.750–4756.500	4827.875–4847.875	Air	mag	(7)	−0.64	1.8
39	aTiO	5445.000–5600.000	5420.000–5442.000	5630.000–5655.000	Air	mag	(7)	−0.74	2.5
40	CaH1	6357.500–6401.750	6342.125–6356.500	6408.500–6429.750	Air	mag	(7)	−0.68	1.9
41	CaH2	6775.000–6900.000	6510.000–6539.250	7017.000–7064.000	Air	mag	(7)	−0.74	2.8
42	NaISDSS[b]	8180.0–8200.0	8143.0–8153.0	8233.0–8244.0	Air	Å	(8)	…	…
43	TiO2SDSS	6189.625–6272.125	6066.625–6141.625	6422.0–6455.0	Air	mag	(8)	−0.79	2.0
44	D4000[c]	…	3750.000–3950.000	4050.000–4250.000	Air	…	(9)	−0.33	1.4
45	Dn4000[c]	…	3850.000–3950.000	4000.000–4100.000	Air	…	(10)	−0.33	1.1
46	TiOCvD[b,d]	…	8835.000–8855.000	8870.000–8890.000	Vacuum	…	(5)	…	…

Notes.
[a] The Ca II triplet and Paschen bands are each meant to be measured as a group using a set of five interspersed continuum passbands, see Cenarro et al. (2001, Section 4.3) for details. However, these indices are currently treated identically to other indices; i.e., independently measured and with two associated sidebands. Future improvements of the `DAP` will provide functionality for these more complex index definitions.
[b] Due to the limited spectral range of the `MILES-HC` library, index does not have a velocity-dispersion correction. These data should be used with caution or ignored.
[c] Bandpass integration is performed over F_ν, not F_λ. The index is defined as the ratio of the red bandpass integral divided by the blue bandpass integral, see Equation (23).
[d] Index defined as the ratio of the blue bandpass integral divided by the red bandpass integral, see Equation (24).
References. (1) Trager et al. (1998), (2) Worthey & Ottaviani (1997), (3) Serven et al. (2005), (4) Cenarro et al. (2001), (5) Conroy & van Dokkum (2012), (6) Spiniello et al. (2012), (7) Spiniello et al. (2014), (8) La Barbera et al. (2013), (9) Bruzual (1983), (10) Balogh et al. (1999).

molecular composition in stellar atmospheres. As listed in Table 4, there are three such indices provided in DR15, the ubiquitous D4000 and Dn4000 and a TiO bandhead. Recall that the D4000 and Dn4000 indices both quantify the strength of the 4000 Å break, but with slightly different definitions for their two sidebands. These three indices have subtly different definitions,[69]

[69] The λ^2 terms in Equation (23) are due to the D(n)4000 indices being defined as an integration in wavelength over flux per unit frequency (f_ν) instead of over f_λ.

as noted in the table; we provide them here for completeness (see Equation (11)):

$$\mathrm{D(n)4000} = \frac{\langle \lambda^2 f_c \rangle_{\mathrm{red}}}{\langle \lambda^2 f_c \rangle_{\mathrm{blue}}}, \tag{23}$$

$$\mathrm{TiO} = \frac{\langle f_c \rangle_{\mathrm{blue}}}{\langle f_c \rangle_{\mathrm{red}}}. \tag{24}$$

As with the definition and use of the emission-line passbands, all spectral-index passbands are defined at rest and then redshifted appropriately for measurements on the *observed* spectra. This leads to the $(1+z)$ factor in Equation (22), such that the indices are always provided in the rest frame. In DR15, we adopt a single redshift, z_0 (see Section 7.1.4), for all spectral-index measurements within a given data cube.

10.1. Velocity Dispersion Corrections

The numerous, blended absorption features in the continuum spectra of these galaxies lead to a dependence of the determination of the pseudo-continuum and the main bandpass integral on the effective resolution of the data. By effective resolution, we mean the convolution of the intrinsic spectrum with the combined kernel made up of the instrumental resolution and the astrophysical Doppler broadening. For comparison of the index measurements from the MaNGA spectra with model grids (e.g., Thomas et al. 2011), it is important that both are performed at the same effective resolution. This is often done by matching the instrumental resolution of the data to a fiducial resolution—such as the original Lick resolution of 8.4 Å—and correcting the measurements to a fiducial astrophysical velocity dispersion.

Instead of degrading the resolution of the MaNGA data, we make the measurements directly at their native resolution. We then calculate a velocity-dispersion correction for each index that converts the measurement to one that would be made if the galaxy had $\sigma_* = 0$. This correction is constructed by calculating the index in both the best-fitting continuum model, $\mathcal{I}_{a,m}$, (from the combined emission-line and stellar continuum fit; see Section 9) and the same model constructed with $\sigma_* = 0$, $\mathcal{I}_{a,m,0}$. The correction depends on the index units such that:

$$\delta\mathcal{I}_a = \begin{cases} \mathcal{I}_{a,m,0}/\mathcal{I}_{a,m}, & \text{for Å units} \\ \mathcal{I}_{a,m,0} - \mathcal{I}_{a,m}, & \text{for magnitude units,} \end{cases} \tag{25}$$

which can be applied to the data to get the corrected index, $\mathcal{I}_a^c$, as follows:

$$\mathcal{I}_a^c = \begin{cases} \mathcal{I}_a \delta\mathcal{I}_a, & \text{for Å units} \\ \mathcal{I}_a + \delta\mathcal{I}_a, & \text{for magnitude units.} \end{cases} \tag{26}$$

Recalling from our discussion of the velocity-dispersion corrections for the stellar kinematics in Section 7.1.5, the "raw" velocity dispersion measurements we make, σ_{obs}, include both the astrophysical stellar velocity dispersion and the difference between the MaNGA and MILES spectral resolution (see Equation (6)). We use this to our advantage in the determination of the velocity-dispersion correction for the spectral indices: by calculating $\mathcal{I}_{a,m}$ based on the best-fit model and $\mathcal{I}_{a,m,0}$ using the templates at their native resolution, we account for both astrophysical dispersion and the difference in resolution between MaNGA and MILES. Therefore, the corrected indices from the DAP can be compared to model grids made for indices measured with $\sigma_* = 0$ at the MILES spectral resolution (e.g., Thomas et al. 2011).

In DR15, we provide $\mathcal{I}_a$ and $\delta\mathcal{I}_a$, such that users can calculate $\mathcal{I}_a^c$ or derive and apply their own corrections. Just as with the stellar velocity dispersion correction, *the user must calculate the corrected spectral indices for themselves*. Importantly, although provided as part of the release, the spectral indices outside of the MILES spectral range should only be used as a rough guide because no velocity-dispersion corrections are available. The velocity-dispersion corrections are critical to both the comparison of the relative absorption-line strengths within and among galaxies, as well as for the inference of stellar-population properties when compared to model values. Spectral indices outside of the MILES spectral range in DR15 will therefore have limited use and should only be used, even in a relative sense, with spectra of similar velocity dispersion.

10.2. Possible Model-driven Biases

Before presenting our quality assessments of the DAP spectral-index measurements, we first comment on two possible sources of systematic error that we do *not* explore in detail here.

First, any systematic error in the emission-line model will propagate to biases in the spectral indices measured with passbands that span emission-line regions because the emission-line model is subtracted from the data before the spectral indices are measured. It is, therefore, important to consult Belfiore et al. (2019) for an in-depth investigation of the systematic errors in the emission-line modeling. Although other sources are possible (e.g., non-Gaussianity in the line profiles), Belfiore et al. (2019) find that the dominant source of systematic error, of those they explored, is the choice of the stellar-continuum templates. Importantly, there is little power in the fit-quality figures-of-merit to discriminate between continuum models based on the MILES-HC library and three different simple stellar-population model template sets. However, based on their comparison between the Bruzual & Charlot (2003, BC03) and MILES-HC templates, Belfiore et al. (2019), show it is possible that the latter yields systematically shallow Balmer absorption-line depths. In the context of the spectral-index measurements, this would imply a systematic under-subtraction of the Balmer emission lines and a systematically small absorption-line equivalent width for Balmer-line indices (e.g., Hβ and Hδ_{A}).

Second, although the raw spectral-index measurements are largely model-independent (apart from the emission-line subtraction), the velocity-dispersion corrections are constructed solely from the stellar-continuum model. Systematic errors due to the velocity-dispersion correction can manifest in two ways: (1) biases due to a propagated bias in the velocity-dispersion measurement itself; and (2) biases due to local inaccuracies in the ability of the model to reproduce the observed spectrum. Effects due to the former are explored in the experiments performed in Section 10.3.1. In the context of the latter, we note discrepancies that Belfiore et al. (2019), find between the Mgb and NaD absorption features between the BC03 and MILES-HC templates (see Figure 8 and Section 4.2 of Belfiore et al. 2019, and see discrepancies in the model fit to the central spectrum of observation 8728-12703 in our Figure 2).[70] In

[70] We note that some of this discrepancy in the NaD absorption depth may be to due interstellar absorption, as opposed to a limitation in stellar parameter coverage.

fact, Figure 4 shows that the range of stellar parameters consolidated into a single MILES-HC template can be quite broad in [Fe/H] and $\log g$. Therefore, it is likely that the MILES-HC library cannot, in detail, mimic the full range of features seen in the full MILES library, which compounds the ability of the entirety of the MILES library to match the detailed features in the MaNGA galaxy spectra due to its limited coverage of stellar parameter space. Nonetheless, the velocity-dispersion corrections are typically a few percent, meaning that the influence of such systematic errors on the corrected spectral index will be of the same order at most.

10.3. Quality Assessments

Similar to the tests of the stellar kinematics in Section 7, we assess the quality of the spectral-index measurements using both idealized simulations (Section 10.3.1) and repeat observations (Section 10.3.2). Although more limited in scope than our stellar kinematics assessments, these tests provide guidance in terms of the limitations of the data provided by the DAP. Specifically, we note that in both binning cases provided in DR15, bins and/or spaxels may be below the S/N needed to meet the goals for a given analysis of the spectral indices. In Section 10.3.3, we discuss the procedure one can use to combine index measurements from multiple bins to improve their precision. A comparison of the results provided by the DAP and the Firefly Value-Added Catalog[71] was performed by Goddard (2018) for the data available as of SDSS DR14 (Abolfathi et al. 2018).[72]

10.3.1. Idealized Simulations

We limit the scope of our idealized simulations to consider measurements of the following subset of indices (see Table 4): Hδ_A, Hβ, Mgb, Fe5270, Fe5335, NaD, and D4000. We also limit the exploration of the systematic and random errors in these indices to their measurement for a subset of single spectra chosen from the MILES-HC templates: template 21, 23, 25, 28, 31, 41, and 45 (see Figure 5). These templates were chosen to roughly span the range of values for the selected indices relevant to the galaxy data. The spectral region of each spectral index (except D4000) are shown in Figure 23 for the subset of seven MILES-HC templates.

We convolve each of the seven selected MILES-HC templates with a wavelength-independent Gaussian with velocity dispersions of 35, 70, 140, 280, and 396 km s^{-1}. No velocity shift is applied, and we also include the original spectrum (i.e., a $\sigma_* = 0$ km s^{-1} spectrum), such that there are six noise-free spectra used for each of the seven MILES-HC templates. Noise is added to each spectrum to meet a g-band S/N of 2^i for $i = 0, \ldots, 7$ with a wavelength dependence that matches the median wavelength dependence of the noise vectors measured for an example MaNGA data cube (7495-12704 in this case). We measure the spectral indices for both the noise-free synthetic spectra and 1000 noise realization for each S/N$_g$.

The "true" velocity-dispersion corrections for each index, unaffected by noise, are calculated using the noise-free spectra and the known input dispersion. However, for each noise realization, the corrections are calculated similarly to the observed galaxy spectra: we first fit the stellar kinematics of each synthetic spectrum using the same code used to fit the galaxy data—including the full MILES-HC library in the fit—to obtain a stellar velocity dispersion that is subject to relevant random and systematic error; this velocity-dispersion measurement is used to construct the correction. With both the true and measured velocity-dispersion corrections, we have explored the accuracy and precision of the velocity-dispersion corrections themselves, as well as their relative influence on the uncertainties in the corrected indices.

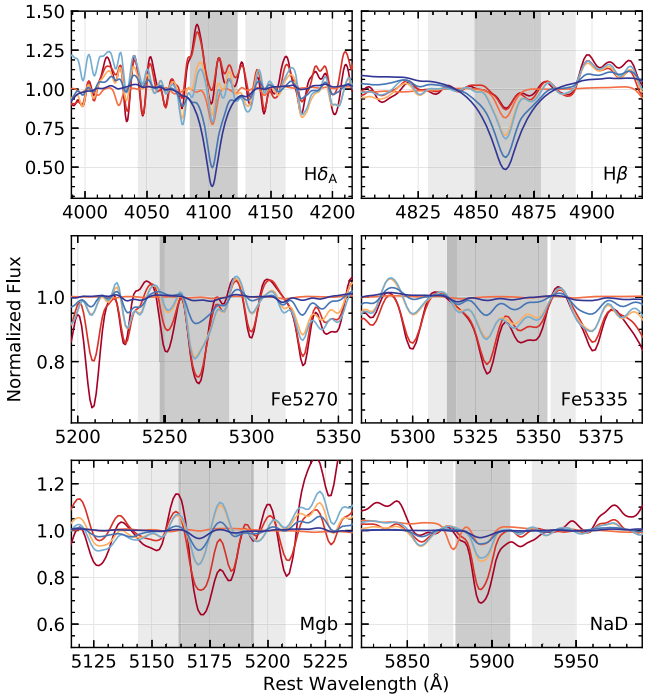

Figure 23. A subset of seven template spectra from the MILES-HC library in six spectral windows near absorption-line indices measured by the DAP. The template spectra have been broadened by $\sigma_* = 140$ km s^{-1}, and their line color is ordered from blue to red by the measured Hβ equivalent width; see Table 5. From top-to-bottom, left-to-right, the spectral regions are near the Hδ_A, Hβ, Fe5270, Fe5335, Mgb, and NaD absorption indices, as marked in the lower right corner of each panel. The dark gray region highlights the main passband of the index, with lighter gray regions showing the two sidebands; note that the sidebands can overlap with the main band (e.g., Fe5335).

For each of the seven spectral indices measured in the idealized simulation, we construct plots like the examples provided in Figure 24. In each column, the top panel shows the accuracy of the mean *corrected* spectral index measured from the simulated spectra, $\langle \mathcal{I}^c_a \rangle$ compared to the known value from the noise-free spectrum, $\mathcal{I}^c_{a,0}$. The middle panel provides the standard deviation in the measured index (i.e., the empirical error), $\delta\mathcal{I}^c_a$, and the bottom panel shows the ratio of $\delta\mathcal{I}^c_a$ to the mean of the error provided by the formal error-propagation calculations performed by the DAP.

The specific indices and templates chosen for Figure 24 are illustrative of the general behavior across the seven indices explored: The accuracy of the measured index depends both on the S/N$_g$ of the spectrum and the intrinsic equivalent width of the feature. Indeed, one should expect that S/N requirements to meet a given *fractional* accuracy are more stringent for weak (shallow) features than for strong (deep) features. Table 5 provides each index measured for each noise-free spectrum and

[71] https://www.sdss.org/dr15/manga/manga-data/manga-firefly-value-added-catalog/

[72] Specifically, see Figure 3.20 at https://researchportal.port.ac.uk/portal/files/11006420/Daniel_Stephen_Goddard_Thesis_ICG_Portsmouth.pdf.

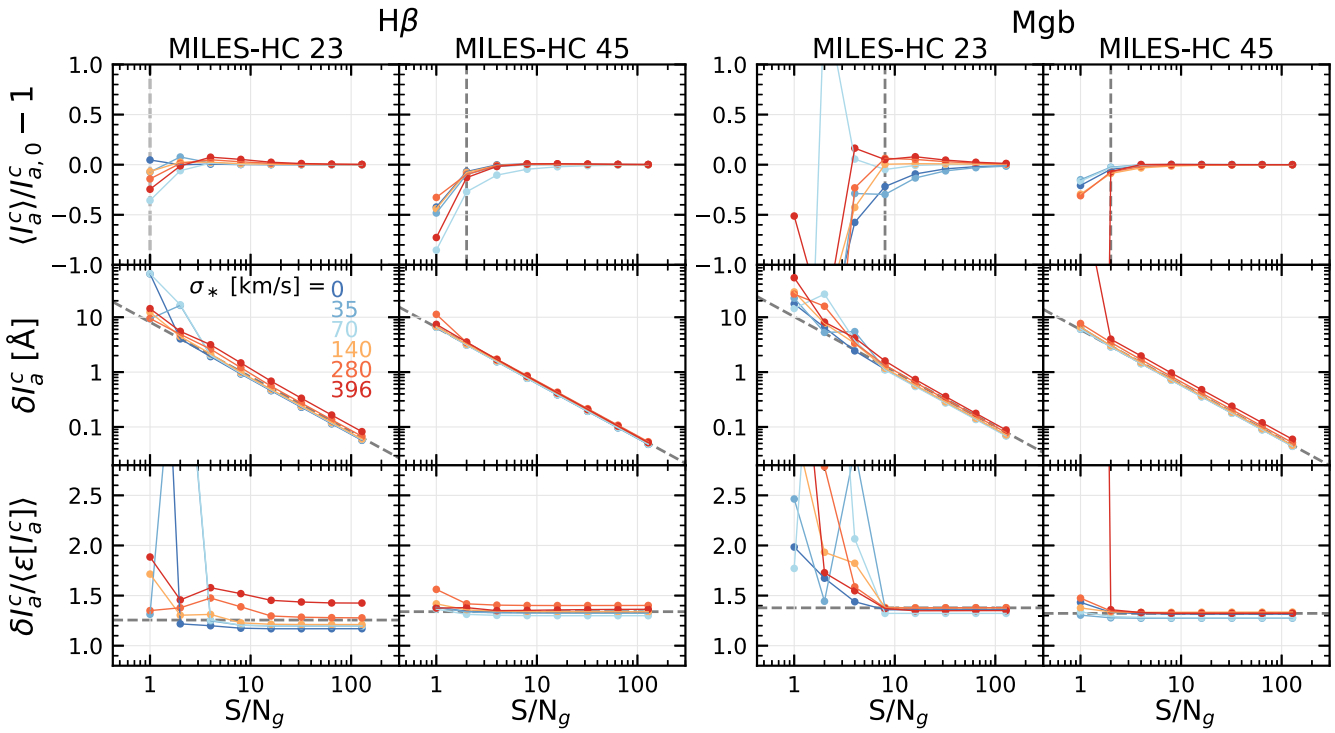

Figure 24. Results for idealized simulations testing the recovery of the Hβ and Mgb absorption-line indices as a function of S/N$_g$ and spectral template. See Section 10.3.1 for a description of the simulations. The results for two MILES-HC templates (template 23 and 45; Figure 5) are shown for each index as marked. Line colors in all panels indicate the stellar velocity dispersion, σ_*, of the LOSVD kernel. Each column of panels shows the fractional recovery of the input index (top), the empirical estimate of the random error (middle), and the ratio of the empirical estimate and formal calculation of the index error. The vertical dashed line in the top panels indicate the S/N$_g$ above which the median and full range of the binned data are less than 30% and 50%, respectively. The dashed line in the middle panels show the optimal inverse proportionality between the index error and S/N$_g$. The dashed line in the bottom panels show the median ratio between the empirical estimate and formal calculation of the error for all S/N$_g$ and σ_* bins.

Table 5
Example MILES-HC Spectral Indices

	HδA						Hβ						Mgb					
ID	(Å)	corr	ε	S/N$_g$	γ	$\delta\epsilon$	(Å)	corr	ε	S/N$_g$	γ	$\delta\epsilon$	(Å)	corr	ε	S/N$_g$	γ	$\delta\epsilon$
28	−5.0	1.049	1.66	16	0.05	1.4	1.2	1.106	0.81	1	0.08	1.3	4.9	0.982	0.73	1	0.08	1.3
45	−6.1	1.030	1.39	4	0.05	1.4	1.2	1.070	0.82	2	0.05	1.3	3.4	0.990	0.77	2	0.02	1.3
31	2.0	1.015	0.81	1	0.05	1.2	1.6	0.999	0.92	4	0.02	1.4	−0.2	1.137	1.15	64	0.05	1.6
41	−1.1	1.058	1.16	16	0.03	1.3	3.3	1.038	0.82	1	0.02	1.3	0.3	0.759	0.84	8	0.07	1.4
21	0.6	0.925	0.95	16	0.51	1.3	3.5	1.032	0.89	1	0.04	1.3	0.2	0.497	0.92	32	0.18	1.5
25	8.0	1.002	0.89	1	0.03	1.2	6.3	1.013	0.87	1	0.17	1.2	0.6	1.046	0.92	8	0.03	1.4
23	11.1	1.003	0.80	1	0.12	1.2	8.2	1.016	0.91	1	0.30	1.3	0.3	1.014	1.01	8	0.04	1.4

	Fe5270						Fe5335						NaD					
ID	(Å)	corr	ε	S/N$_g$	γ	$\delta\epsilon$	(Å)	corr	ε	S/N$_g$	γ	$\delta\epsilon$	(Å)	corr	ε	S/N$_g$	γ	$\delta\epsilon$
28	4.2	1.087	0.70	1	0.03	1.3	4.5	1.103	0.68	1	0.02	1.5	4.8	1.032	0.74	1	0.03	1.2
45	3.8	1.084	0.80	1	0.01	1.4	3.7	1.114	0.80	1	0.03	1.5	3.3	1.045	0.86	1	0.04	1.2
31	0.1	0.912	1.08	128	0.08	1.4	0.0	1.223	1.46	128	0.05	1.9	0.6	0.830	1.37	16	0.09	1.3
41	2.9	1.075	0.91	1	0.07	1.4	2.6	1.147	0.94	2	0.02	1.6	1.5	1.062	1.06	2	0.01	1.2
21	2.8	1.073	0.96	2	0.05	1.3	2.8	1.142	1.04	2	0.05	1.5	1.3	1.059	1.19	8	0.04	1.2
25	1.1	1.044	1.00	2	0.06	1.4	1.0	1.150	1.09	8	0.02	1.6	0.7	1.040	1.23	8	0.05	1.3
23	0.1	0.959	1.08	16	0.07	1.4	0.2	1.147	1.33	16	0.06	1.6	0.4	0.991	1.35	16	0.07	1.3

Note. Six columns are provided for each index: (1) the uncorrected index value for $\sigma_* = 140$ km s^{-1}, (2) the index correction for $\sigma_* = 140$ km s^{-1}, (3) the normalization of the inverse correlation between the empirical error and S/N$_g$, (4) the S/N$_g$ below which the fractional systematic error is significant, (5) the maximum ratio between the systematic and random error for all S/N$_g$ and σ_* bins, and (6) the mean ratio between the empirical error and the DAP-calculated formal error. See the text of Section 10.3.1 for additional description.

shows that the corrected Hβ index for template 23 (8.3 Å) is much larger than for template 45 (1.3 Å), whereas the opposite is true of the Mgb index. Thus, we find that the S/N$_g$ requirements for a fixed fractional accuracy of Hβ are lower for an early-type (roughly A-type) star than for a late-type (roughly K-type) star (see Figure 5), and vice versa for the Mgb index.

We crudely quantify this effect by determining the S/N_g above which the median and full-spread in index recovery for all σ_* bins are better than 30% and 50%, respectively. This "minimum" S/N_g is marked as a vertical dashed line in the top panels of Figure 24, and provided for all templates and indices used for this test in Table 5. We note that, although it is generally true that the S/N_g limit is higher for weak features, the detailed correlation is non-trivial.

The middle panel of each column in Figure 24 shows the tight inverse correlation between the random error in each *corrected* spectral index and the S/N_g. This correlation has relatively weak secondary dependencies on σ_* and the value of the spectral index. We characterize the primary relation as an inverse proportionality and determine the optimal proportionality constant such that $\delta\mathcal{I}^c_a = 10^{\varepsilon}(S/N_g)^{-1}$. We note that this relation is generally well-posed and robust, whereas a similar relation use to quantify the *fractional* error (i.e., $\delta\mathcal{I}^c_a/\langle\mathcal{I}^c_a\rangle$) is not, particularly at low equivalent width. The proportionality constants, 10^{ε}, are used to construct the dashed lines in the middle panels of Figure 24, and ε is provided for all spectral indices and templates in Table 5. For example, the coefficients in Table 5 yield a typical random error in the corrected Hβ index at $S/N_g = 10$ of 0.65–0.83 Å, depending on the spectrum.

For small values of $\mathcal{I}^c_{a,0}$, the fractional systematic error is not well posed and it is more informative to consider the ratio of the systematic and random error, $\gamma = (\langle\mathcal{I}^c_a\rangle - \mathcal{I}^c_{a,0})/\delta\mathcal{I}^c_a$. In *all* cases (independent of index, template, velocity dispersion, or S/N_g), the systematic error from these idealized simulations is less than the random error, $\gamma < 1$, most often substantially so. In addition, we do not find any significant trends in this ratio with S/N_g or σ_*. For reference, we provide the *maximum* value of γ for any (S/N_g, σ_*) bin for all indices and templates in Table 5. We note that the largest γ values occur at high S/N_g, where the random error is small (e.g., the Hδ_A index for template 21).

The contribution of the velocity-dispersion correction to the total error in $\mathcal{I}^c_a$ is relatively small, despite what can be substantial error in the correction itself toward low S/N_g (e.g., $5\% \pm 5\%$). The primary reason is that the corrections are generally only a few percent such that the fractional error in the correction itself (e.g., 0.05/1.05) is small, and therefore so is its contribution to the total error. This becomes less true toward low S/N_g, where the corrections can suffer from substantial systematic error (e.g., as seen in the Mgb index for template 23 and $S/N_g \lesssim 10$). We also note a correlation between the error in the index correction and σ_*, which leads to the increase in $\delta\mathcal{I}^c_a$ with σ_* for Hβ in template 23 seen in Figure 24; however, we do not provide a general usage recommendation because this effect has widely varying influence on any given index for any given underlying spectrum.

The bottom panel of each column in Figure 24 shows the ratio between the empirically measured spectral-index error, $\delta\mathcal{I}^c_a$, and the mean of the formally propagated errors provided by the DAP, $\langle\epsilon[\mathcal{I}^c_a]\rangle$. Generally speaking, the true random errors in the spectral indices are larger than those provided by the DAP. We note, in particular, that the formal calculations in the DAP do not include any error from the velocity-dispersion correction (no error is calculated on the correction). Although there can be systematic differences in the accuracy of the formal errors between spectra of different σ_* and at different S/N_g (e.g., the Mgb index in MILES-HC template 23), there is little motivation to try to capture these variations here and we simply determine the median value of the ratio, provided as $\delta\epsilon$ in Table 5. We return to the accuracy of the DAP-provided formal errors using the repeat observations.

We have excluded the results for D4000 from Table 5, summarizing them here instead. In contrast to the results from the absorption-line indices, the formally propagated errors for D4000 are accurate such that we always find $\delta\epsilon = 1$. Given the difference between the calculation of the bandhead and absorption-line indices, we infer that part of the inadequacy of the formal error calculations for the absorption-line indices is due to an insufficient propagation of the error introduced by the continuum calculation (see Equation (11); nor does it properly propagate the error in the index correction, as we mentioned above). Although we calculate velocity-dispersion corrections for both the bandhead and absorption-line indices, these corrections are less important for the former; the velocity-dispersion correction for D4000 measured for all templates used in our idealized simulations are less than 0.1% for $\sigma_* = 140$ km s^{-1}. Of the seven templates included, only the late-type templates 28 and 45 show any systematic error in D4000; these extremely red spectra (D4000 $\gtrsim$ 3.5) have a "minimum" $S/N_g \sim 4$. Finally, the proportionality constants for the trend of random error with S/N_g in D4000 are generally smaller than for the absorption-line indices, with errors of a few hundredths at $S/N_g = 10$ for early-type spectra; for late-type spectra, these errors reach a few tenths at the same S/N_g.

10.3.2. Repeat Observations

Using the repeat observations of 56 galaxies (Table 1), we assess the random uncertainties in the spectral-index measurements and the accuracy of the DAP error calculations. The spatial registration of the repeated observations is discussed in Section 7.5.1.

Figure 25 shows the results for six of the measured indices provided in DR15: from left-to-right and top-to-bottom, Hδ_A, D4000, Hβ, Fe5335, Mgb, and NaD. The mean and difference in the spectral indices measured for paired spaxels between repeated observations are calculated and binned by their S/N_g and stellar velocity dispersion, σ_*.

Figure 25 provides a set of three panels for each of the six spectral indices. The top panel of each group shows the distribution of the index measurements in (S/N_g, σ_*) bins using a box-and-whisker plot. For each bin, the central box spans the inner two quartiles of the spectral-index distribution and the lines span its full extent. These distributions are a result of the convolution of the error distribution with the intrinsic, astrophysical distribution. The latter is evident in, for example, the higher D4000 measurements for galaxy regions with larger velocity dispersion. The middle panel shows the empirical estimates of the random error computed by taking half of the 68% confidence interval and dividing by $\sqrt{2}$. Finally, the bottom panel gives the ratio of the empirical error to the median of the DAP-provided error.

Similar to the idealized simulations in the previous section, we find that an empirical determination of the error is well-characterized by an inverse proportionality with S/N_g. The trends with S/N_g show stronger deviation from the optimized relation when compared to Figure 24 for both Hβ and Mgb; however, this is expected given the variation in Table 5 of ε of a few tenths between different templates. In other words, we have not differentiated between the spectra included in the

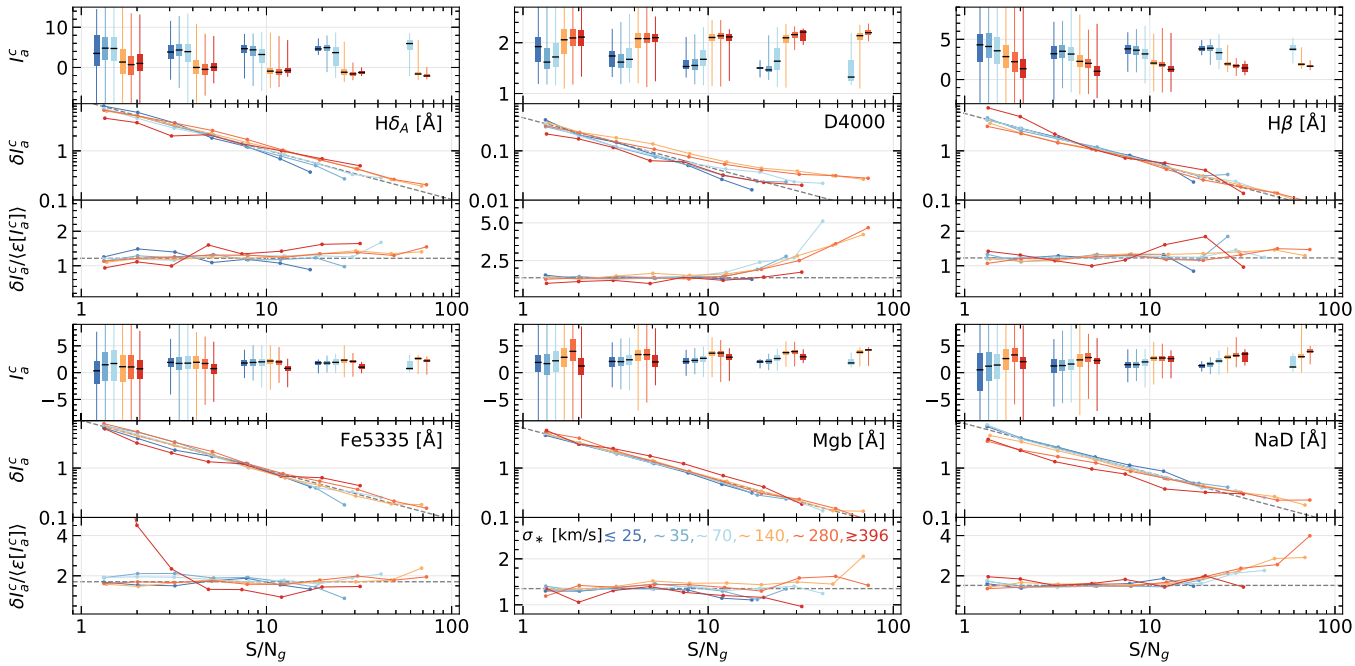

Figure 25. Assessments of the random errors in six spectral indices determined from repeat observations of 56 galaxies. Three panels are provided for each index: the top panel gives box-and-whisker representations of the distribution of index measurements in bins of S/N_g and σ_*; the σ_* bins are colored as labeled in the bottom panel of the Mgb panel group. The box spans the inner two quartiles of the distribution, and the whiskers span its full extent. The middle panel in each group provides the empirical estimate of the error in the index value for each (S/N_g, σ_*) bin. The dashed-gray line is the optimized trend for all (S/N_g, σ_*) bins assuming an inverse proportionality between the error and S/N_g. The bottom panel shows the ratio of the empirically estimated error to the mean error provided by the formal calculations in the DAP. The dashed-gray line shows the median ratio across all (S/N_g, σ_*) bins.

(S/N_g, σ_*) bin (e.g., by also binning by index value) such that this variation in ε is reflective of the astrophysical variation in the spectra falling into each bin.

We also find that the errors reported by the DAP tend to underestimate the empirically estimated errors by a ratio that is roughly constant at $S/N_g \lesssim 15$. At higher S/N_g, D4000 and NaD, in particular, show an increase in the difference between the empirical and formal estimates of the error, similar to the behavior seen for the stellar kinematics in Figure 19 (and the emission-line properties explored by Belfiore et al. 2019). These difference in the repeated observations are due to astrometric errors in the dithered MaNGA observations. We expect that this increase is *not* seen in the empirical error of the other indices in Figure 25 because the random errors in the measurement of the index itself dominate over the contribution from the astrometric errors.

In Table 5, we provide measurements of both ε and $\delta\epsilon$ for all indices measured within the MILES spectral range; i.e., those indices where we have been able to calculate velocity-dispersion corrections. For $\delta\epsilon$, we only consider data with $S/N_g < 15$. We find that the errors are typically ~0.5–1 Å or $\lesssim$0.03 dex at $S/N_g \sim 10$ for indices with units of Å or magnitudes, respectively. Finally, we find that the errors reported by the DAP are smaller than the empirically estimated errors by ~30%–100%.

10.3.3. Binning Spectral Indices

In DR15, we provide spectral indices measured both for individual spaxels, as a product of the hybrid-binning approach, and for the spectra resulting from Voronoi-binning the data to $S/N_g \gtrsim 10$. However, the precision of the spectral-index measurements for a given science pursuit can require measurements at significantly higher S/N (e.g., Parikh et al. 2019). Instead of providing additional DAP output resulting from binning to a higher S/N_g threshold, we explore and test the accuracy of a method of combining the indices directly.

Given the use of a linear continuum (Equation (11)), the explicit calculations of the absorption-line indices from Equation (22) are, to good approximation,

$$\mathcal{I}_a \approx \begin{cases} \dfrac{1}{1+z}[1 - S(f_c)/S(C)], & \text{for Å units} \\ -2.5\log[S(f_c)/S(C)], & \text{for magnitude units.} \end{cases} \quad (27)$$

For Å units, these definitions can be used to derive

$$\mathcal{I}'_a[\text{Å}] = \frac{1}{1+z}\left[1 - \frac{\sum_i S(f_{c,i})}{\sum_i S(C_i)}\right] = \frac{\sum_i S(C_i)\mathcal{I}_{a,i}}{\sum_i S(C_i)}, \quad (28)$$

where $\mathcal{I}'_a$ is the spectral index measured for a spectrum constructed as the sum of $i = 0..N-1$ spectra with spectral indices $\mathcal{I}_{a,i}$. In other words, the combined spectral index is the weighted sum of the indices from the individual spectra, where the weights are the value of the continuum integrated over the main passband. Equivalently for magnitude units, we find

$$\begin{aligned} \mathcal{I}'_a[\text{mag}] &= -2.5\log\left[\frac{\sum_i S(f_{c,i})}{\sum_i S(C_i)}\right] \\ &= -2.5\log\left[\frac{\sum_i S(C_i)10^{-0.4\mathcal{I}_{a,i}}}{\sum_i S(C_i)}\right]. \end{aligned} \quad (29)$$

In principle, these derivations suggest that a simple linear combination of the spectral indices can be used to construct the

measurement in a summed spectrum. We test this result directly as follows. Selecting the 112 PLATEIFUs from Table 1 that compose the first two observations of each target, we run the DAP four times, once using the hybrid-binning approach and three times using a Voronoi-binning approach that adopts thresholds of $S/N_g = 10$, 20, and 40, respectively. For spectra composed of more than one spaxel and directly analyzed in the three Voronoi-binning cases, we construct the combined index using Equation (28) using the individual spaxel measurements from the hybrid-binning approach and compare those to the measurements made directly using the binned spectrum. In this experiment, we apply Equation (28) *after* correcting the spectral indices for the velocity dispersion and we replace the integral of the continuum over the main passband by the mean *g*-band flux. The latter is to make this experiment most directly applicable for users of the DR15 data; the calculation of $S(C)$ is not provided in DR15, whereas the mean *g*-band flux is provided. Calculations of $S(C)$ will be provided in future releases to ensure a more accurate calculation. The results are shown for six spectral indices—Hβ, Hδ_A, D4000, Fe5335, Mgb, and NaD—in Figure 26.

Figure 26 provides two panels for each of the six spectral indices. The left panel shows the distribution of the difference in the *corrected* spectral indices as a function of the S/N_g in the binned spectrum. We overplot the calibrated error relations from Table 4, as well as lines representing factors of 10 and 100 below this relation. We calculate, and provide in the figure, the percentage of points with differences that are larger than the expected error, between each of the rescaled relations, and below 1% of the expected error. The second panel provides the distribution of the difference normalized by the expected error and marginalized over all S/N_g. We calculate, and provide in the figure, the mean and standard deviation of the distribution. With the exception of D4000, the difference between a direct spectral-index measurement using a binned spectrum is consistent with the results of Equation (28) to better than 10% of the expected error for the majority of the calculations and with a systematic shift of at most a few percent of the expected error.

Depending on the accuracy and precision needed, this approach of combining indices could be used directly for scientific inquiry, or at least provide guidance for follow-up measurements. We emphasize that this is a *statistical* statement; i.e., systematic errors in the combined index compare to a binned spectrum will largely average out of a study over many galaxies. Conversely, studies focused on small regions of individual galaxies should always measure the indices directly on the binned spectra to ensure the measurement accuracy.

10.3.4. Summary

In our assessments of the quality of the spectral indices provided as part of DR15, we find:

1. Our idealized simulations show that systematic errors can be significant with respect to the value of the index, particularly at low S/N and low EW. However, the systematic error is *always* less than, and typically less than 10% of, the random error at any S/N_g. For the subset of spectral indices investigated, we provide the maximum expected systematic error as a fraction of the random error (γ) in Table 5.
2. Random errors, determined using both idealized simulations and repeat observations, are well behaved to very low S/N ($S/N_g \gtrsim 2$), following a simple inverse proportionality with S/N_g. Using repeat observations, we calibrate this relation for all spectral indices provided in DR15 such that users can determine the S/N_g required to meet a desired spectral-index error using the data (ε) in Table 4.
3. The random errors reported by the DAP are underestimated, as determined for both the idealized simulations and the repeat observations. For all the indices in DR15, we provide a simple scaling of the reported errors ($\delta\epsilon$ in Table 4) to match the results from the repeat observations. However, in spectra with $S/N_g \gtrsim 15$, the spectral-index measurements can be affected by the astrometric errors in the registration of the dithered observations, as evidenced by stronger discrepancies between repeat observations.
4. We provide a method that combines the index measurements from multiple spectra, which avoids having to recompute the index on the binned spectra themselves. This method yields results that are typically consistent with a direct measurement to better than 10% of the calibrated error (ε from Table 4). However, this is a *statistical* statement that should be treated with caution, or ignored, when applied to a limited number of spatial regions.

10.4. Flagging

The flags used for the spectral-index measurements are virtually identical to those used for the emission-line moments (Section 9.7.1; see Table 8). The NOVALUE, UNRELIABLE, and MATHERROR flags have the same meaning. However, any regions without an emission-line model subtraction are not flagged as NOCORRECTION in a synonymous way to the model subtraction performed for the emission-line moments. Instead, measurements are flagged as NOCORRECTION if there was an error in the calculation of the velocity dispersion correction, or if one could not be calculated because of the spectral range of the fitted models. The latter is a critical consideration for the spectral indices provided in DR15 with passbands at $\lambda \gtrsim 7400$ Å.

11. Performance

In this section, we assess the *overall* performance of the DAP. The performance, specifically with regard to the stellar kinematics, is discussed in Section 7.5; it is discussed with regard to the emission-line modeling in our companion paper, Belfiore et al. (2019); and it is discussed with regard to the spectral indices in Section 10.3. Here, we start with basic statements concerning the success rate of the DAP (Section 11.1) and we then provide a more detailed look at the statistical performance of the two full-spectrum-fitting modules (Section 11.2). In the latter, we note particular regimes where the DAP performs poorly (Section 11.2.2), which will become a focus for future development.

11.1. Success Rate

Although not necessarily relevant to the quality of the data that it provides, the DAP executes successfully for the vast

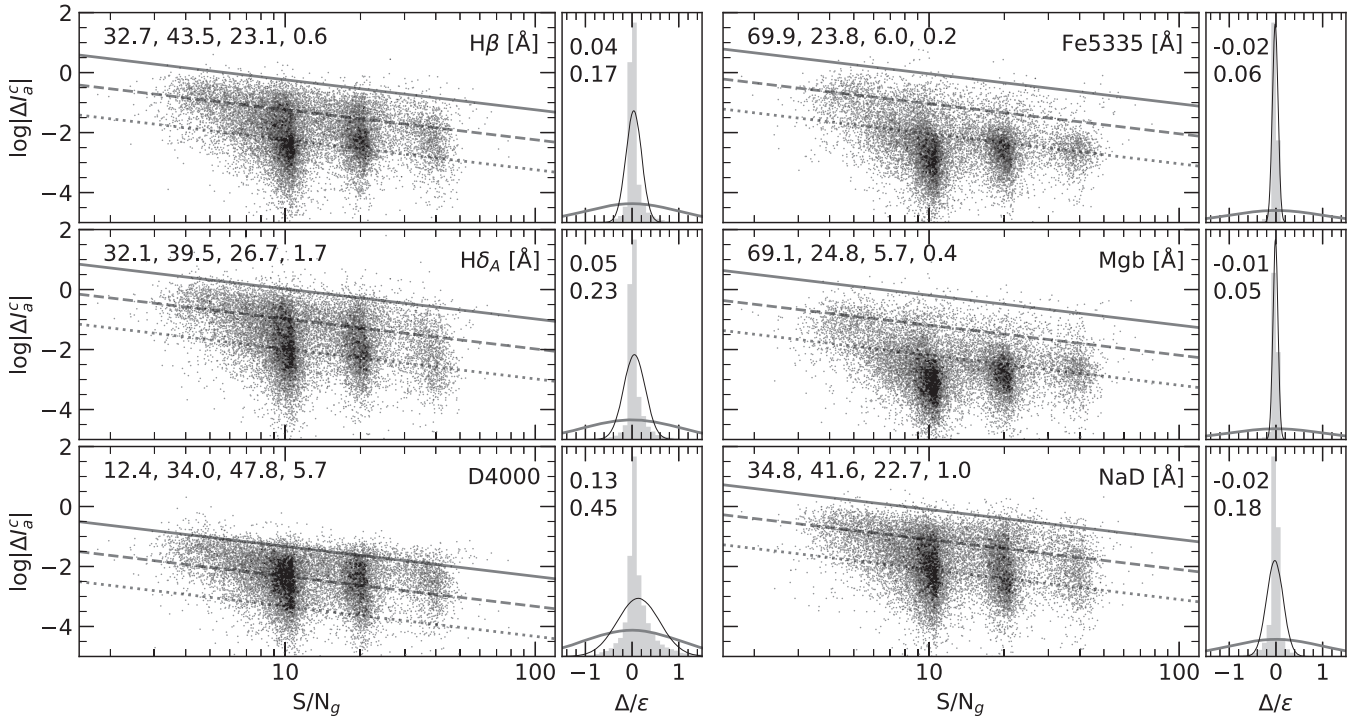

Figure 26. A comparison of six spectral indices determined by direct measurements on a binned spectrum to the result of combining measurements from individual spectra in the bin. See the description in Section 10.3.3. For each index, we show the difference in the two measurements, $\Delta\mathcal{I}_a^c$, as a function of S/N_g, and we show the distribution of ratio of the difference to the error, Δ/ε, over all S/N_g. Gray lines underlying the distribution of $\Delta\mathcal{I}_a^c$ are based on the expected error relation derived in Section 10.3.2; the solid line is the nominal relation, whereas the dashed and dotted lines show, respectively, a factor of 10 and 100 below the nominal error. In these panels, the relevant index is given in the upper right corner, and the numbers in the upper left provide, from left-to-right, the percentage of measurements below the dotted line, between the dotted and dashed lines, between the dashed and solid lines, and above the solid line. For example, 43.5% of the Hβ measurements show differences that are between 1% and 10% of the nominal error in a single measurement. The distributions of Δ/ε are shown in gray, with the Gaussian distribution based on the nominal error relation show in dark gray. In the upper left corner, we provide, from top-to-bottom, the mean and standard deviation of the distribution after clipping 10σ outliers; the relevant Gaussian distribution is plotted as a thin black line.

majority of the data cubes provided in DR15. However, there are still some corner cases where the DAP ends in error for reasons that we are still investigating. For the 4731 data cubes that the DAP attempted to analyze in DR15 (Section 4) using two different methods (DAPTYPEs; Sections 4 and 12), 22 of the 9462 (0.2%) executions failed. The observations with DAP failures are: 7443-3703, 8140-6101, 8146-3702, 8158-3703, 8309-3703, 8312-6101, 8481-6103, 8549-12703, 8993-1901, 9025-12702, 9507-12702, 9677-12703, 9888-9102. Observations 8481-6103, 8549-12703, 8993-1901, and 9507-12702 were successful for the VOR10-GAU-MILESHC method but failed the HYB10-GAU-MILESHC approach. In these cases, the successful VOR10-GAU-MILESHC results are provided, and one can select those galaxies that were successfully analyzed using *both* approaches using the DAPDONE and DAPTYPE columns in the DAPall catalog (Table 11).

11.2. Full-spectrum Fitting

Much of the data provided by the DAP are the result of its two full-spectrum-fitting modules, as described in Sections 7 and 9. The question we address here is: *How well does the* DAP *model each MaNGA spectrum?*

11.2.1. Fit Quality

For each model, m_i, fit to each MaNGA flux measurement, f_i, the DAPcalculates the absolute value of the residual ($|\Delta_i| = |f_i - m_i|$), the fractional residual ($|\Delta_i|/m_i$), and the error-normalized residual ($|\Delta_i|/\epsilon_i$) for each spectral channel, i. We consider the growth of these quantities over each fitted spectrum and the following reduced metrics over the full spectrum: (1) the root-mean-square (rms) of the fit residuals, (2) the rms of the fractional residuals (fRMS), and (3) the χ^2 statistic (the sum of the square of the error-normalized residuals). For assessments of the latter, we use the reduced χ^2, $\chi^2_\nu = \chi^2/(N-\nu)$, where N is the number of fitted spectral pixels and ν is the number of fitted parameters. For the stellar-continuum module, ν is the sum of the number of kinematic parameters (2), the order of the additive polynomial (8), and the number of templates with non-zero weight. For the emission-line module, ν is the sum of the order of the multiplicative polynomial (8) and the combination of the number of templates with non-zero weight and the relevant number of free kinematic parameters associated with those templates (i.e., if a line is not given any weight, the kinematics parameters associated with only that line are not included in ν). These metrics are calculated for (i) the stellar-continuum fit used to determine the stellar kinematics (Section 7), (ii) the combined emission-line and stellar-continuum fit used to determine the emission-line properties (Section 9), and (iii), except for the growth metrics, in 15-pixel regions around each emission line (see Belfiore et al. 2019).[73]

[73] The metrics used in this paper have been recalculated post DR15 given some minor errors in their calculation in DAP version 2.2.1. Code that can be used to recalculate these metrics given the data provided in DR15 is given in the DAP github repository found at https://github.com/sdss/mangadap.

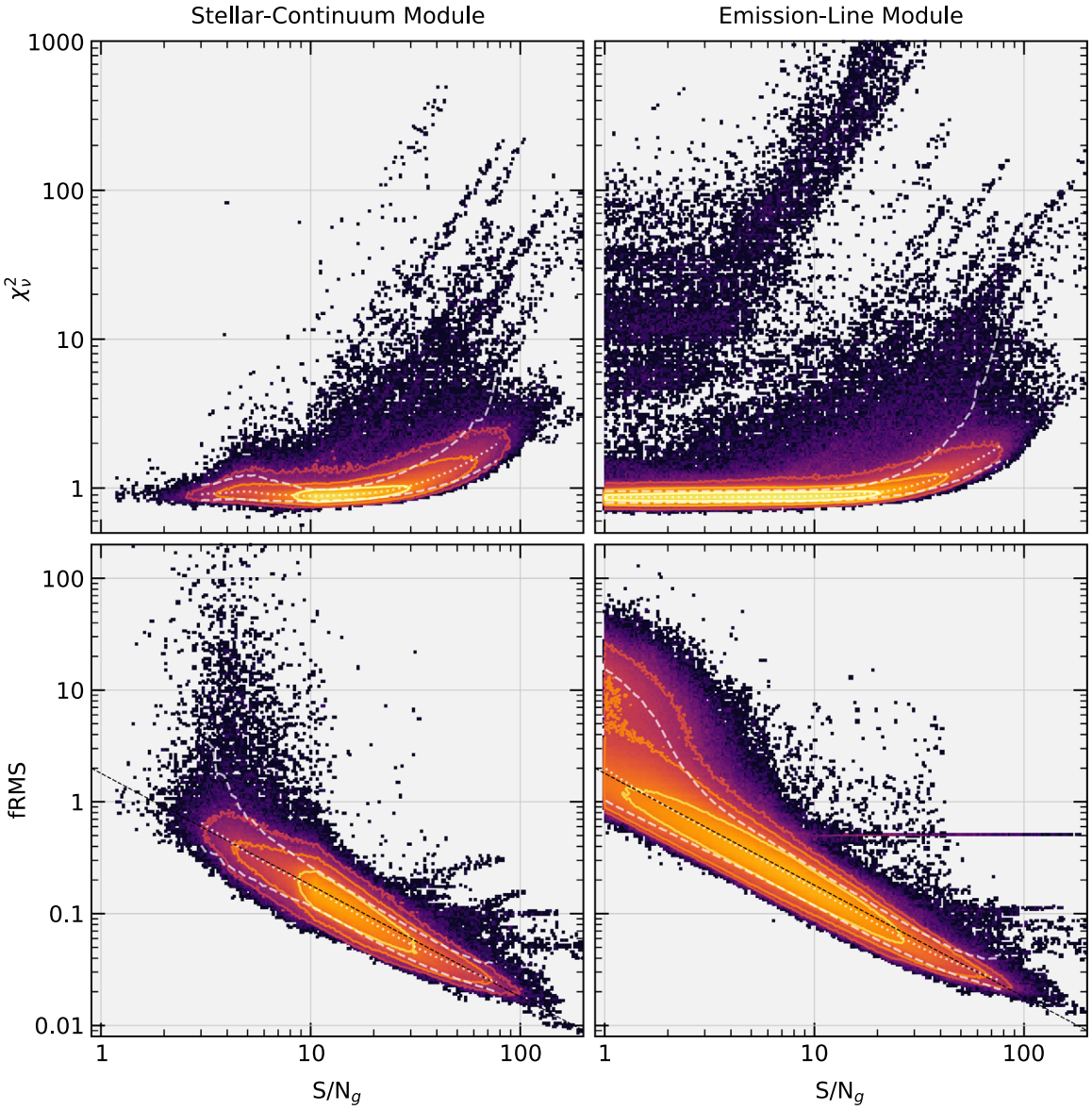

Figure 27. The distribution of χ^2_ν (top) and fRMS (bottom) as a function of S/N$_g$ for all spectra fit in DR15 using the HYB10-GAU-MILESHC approach, excluding data cubes marked as CRITICAL by the DRP or results masked by the relevant DAP module. The difference in the S/N$_g$ distributions for the results of the stellar-continuum module (left) and the emission-line module (right) arises because, for the hybrid approach, the stellar-continuum module analyzes the Voronoi-binned spectra, whereas the emission-line module uses the individual spaxels. Note that the S/N$_g$ distribution in the left panels show there are quite a few binned spectra that do not meet the S/N$_g \gtrsim 10$ threshold. The density of spectra at each location is indicated by the color, where density increases from darker to lighter colors. The three colored contours in each panel enclose 68%, 95%, and 99% of the fitted spectra. That is, populated regions outside the largest contour represent 1% of all MaNGA spectra fit in DR15. The dotted and dashed white lines show, respectively, the median and 95% interval at fixed S/N$_g$.

For models that are well-fitted to the data and assuming robust DRP-provided flux errors, (1) RMS should be proportional to the noise in each spectrum, (2) fRMS should be inversely proportional to its S/N$_g$, and (3) χ^2_ν should be very close to unity (i.e., the mean value of $|\Delta|/\epsilon$ should be nearly unity). Figure 27 demonstrates that these expectations are well met for both fRMS and χ^2_ν resulting from both the stellar-continuum and emission-line fitting modules.

Toward low S/N$_g$, the median χ^2_ν is ∼0.9, meaning that either the model is slightly over-fitting the data (e.g., by an error-driven selection of templates that are a marginally better fit the data) or, more likely, that the errors in the data are slightly overestimated (by approximately 5%). For S/N$_g \lesssim 5$, there are a number of fits with quite large fRMS that are the result of fits to spectra with very low flux levels. This indicates that there may be a systematic underestimation of the continuum level in these S/N$_g$ regimes. Toward high S/N$_g$, the sharp lower limit in χ^2_ν seen at all S/N$_g$ increases from $\chi^2_\nu \sim 0.7$ at S/N$_g \sim 1$ to $\chi^2_\nu \sim 2$ by S/N$_g \sim 100$. This is expected and is a result of the systematic model errors gradually beginning to dominate over the random errors in the observations (see Section 7.5; Belfiore et al. 2019). However, the width of the χ^2_ν distribution increases dramatically toward high S/N$_g$ due to a combination of the paucity of spectra at such high S/N$_g$ and the tendency of spectra that are poorly handled by the DAP (see below) to have high S/N$_g$.

11.2.2. Figure-of-merit Outliers

Figure 27 shows contours that enclose 68%, 95%, and 99% of the fitted spectra for the purpose of highlighting that strong outliers, particularly for the χ^2_ν distribution, or poor fits represent fewer than 1% of all the spectra fit by the DAP. Coherent structures exist for both χ^2_ν and fRMS, such as the data groupings with roughly linear correlations between S/N_g and χ^2_ν and roughly constant values of fRMS. Most of these groupings come from a small number of individual data cubes and they typically fall into one of the following categories:

1. *Unmasked foreground stars*: Although many of the foreground stars that land in the FOV of each IFU are masked by the DRP, this masking is incomplete. Masks for the foreground stars were initially constructed thanks to a by-eye inspection performed by L. Lin and K. Masters. Recently, we have crowd-sourced this inspection using the Galaxy Zoo: 3D interface;[74] objects identified as stars by at least 10 people in that project are masked at present. The incompleteness of the masking is partly by design to avoid accidentally flagging point-like components of the targeted galaxy (e.g., H II regions). Obviously, this results in some stars being missed. In other cases, the foreground star is masked, but the masked area is too small to capture the wings of the stellar PSF. Since the DAP assumes every spectrum in each data cube is of the primary target, these interloping foreground stars will be poorly fitted. Because of their varying luminosity, these outliers can occur over a large range in S/N_g, but they are typically identified as significant χ^2_ν outliers at fixed S/N_g. One should expect these to be outliers in both the stellar-continuum and emission-line modules. Most of the outliers with roughly constant fRMS at high S/N_g, or which roughly follow $\chi_\nu \propto S/N_g$, are due to these unmasked stellar spectra.
2. *Non-targeted Galaxies in the Field-of-View*: Many MaNGA observations include multiple objects in the FOV, which may or may not have been previously recognized as objects superimposed along the line of sight. There are two ways that these objects can lead to poor DAP fits. First, if the objects overlap in a given spectrum, then the DAP will tend to optimize the fit to the more luminous component and a poor fit is likely, depending on the surface-brightness ratio and velocity separation of the two objects. These cases can be difficult to identify because the effect on χ^2_ν can be subtle. However, in the second case, strong deviations in χ^2_ν will occur for regions dominated by the interloper when it is outside of the redshift range allowed by each fit ($\pm$2000 km s^{-1} from the input redshift, z_0, typically from the NSA; Section 7). Poor fits in this category occur at all S/N_g and can typically be identified by the deviation from the χ^2_ν distribution of good fits at similar S/N_g.
3. *Bright/Broad Emission Lines*: The DAP currently assumes that all emission lines are single Gaussian components. However, particularly for very bright emission lines, a second broad component is apparent in the data that cannot be reproduced by the nominal DAP model. Recall that the stellar-continuum fit uses a fixed $\pm$750 km s^{-1} mask for the emission lines offset by the input redshift (z_0; Figure 11). This is typically sufficient to mask the relevant velocity range of the emission features, but it is not sufficient for sources with broad-line regions (e.g., AGN). Both effects can lead to dramatic χ^2_ν outliers from both the stellar-continuum and emission-line fitting modules. Because the latter optimizes the combined continuum+emission-line spectrum, the fit and resulting χ^2_ν are generally better than for the stellar-continuum module; however, the poor quality persists for the brightest/broadest spectra. *Any AGN-focused studies should be very careful with the data provided by the* DAP. *It is likely they will require other analysis products that better handle broad, multi-component emission lines.* In principle, including additional emission-line components in the DAP is straightforward, with minimal code development; however, the validation and stability of the approach will likely require a significant investment.
4. *Unmasked Cosmic Rays*: The DRP removes the vast majority of the cosmic rays from each MaNGA exposure; however, some are still missed for a variety of reasons. Most often the affected spectral channels are easily identified as 3σ outliers and rejected during the fit iterations for both the stellar-continuum and emission-line modules. However, in the emission-line module, these cosmic rays may not be rejected if they fall close enough to emission lines: To avoid rejecting emission-line flux for particularly strong lines, we do not allow the rejection iteration to remove pixels near a fitted emission-line.[75] Lingering cosmic-ray artifacts generally show up as "beam-sized" regions with large values in maps of χ^2_ν. We also note that, for the subset of spectra that we have inspected directly, spectra that populate the upper-left of the upper-right panel in Figure 27 are in fact due to cosmic rays near emission-lines. It is therefore reasonable that a similar distribution is not seen for the results of the stellar-continuum module. Because the cosmic rays affect such a small portion of the spectrum, the model may actually be perfectly reasonable for these cases. Unfortunately, this needs to be assessed on a case-by-case basis. In particular, we have found cases (e.g., 8319-3704) where cosmic rays have dramatically affected the ionized-gas velocity by pulling all lines so that a single line can better fit the cosmic ray.

Beyond these four categories, there are other more subtle limitations of the full-spectrum-fitting modules when one isolates fits that are only just outside the main χ^2_ν distribution. Some interesting examples include galaxies with significantly asymmetric emission-line profiles and star-forming galaxies with easily identifiable emission lines that are not currently in the list of lines fit by the DAP (see Table 3 and Figure 2). We continue to identify these astrophysically interesting phenomena that stretch beyond the standard DAP assumptions and work toward improvements that can properly handle the large variety of MaNGA spectra.

[74] A citizen science project at https://www.zooniverse.org/projects/klmasters/galaxy-zoo-3d.

[75] Specifically, pixels are excluded from rejection if the best-fit emission-line model has a flux density of $>10^{-23}$ erg s^{-1} cm^{-2} Å^{-1} spaxel^{-1}.

12. Output Products

We have touched on the output products provided by the DAP throughout our paper, particularly when introducing some salient details about the data in Section 2 and when discussing the DAP workflow in Section 4. In the latter, we noted that each of the six main DAP modules produces a reference file, which includes all the data produced by the module and can be used to reconstruct the state of the relevant Python object (see Figure 3) to minimize redundant analysis steps. The final step of the DAP is to consolidate and reformat the data in these reference files into the two main files meant for general use, the MAPS file (Section 12.1) and model LOGCUBE file (Section 12.2).

Reformatting the data is a key component of this final step. Most of the core functionality of the DAP treats each spectrum independently, regardless of whether it is from a bin or an individual spaxel. The format of the reference files matches this structure, with spectra organized along rows of 2D arrays and derived quantities organized in data tables with one row per spectrum. However, to ease its use, we provide the data in a spatial format that exactly matches the DRP-produced data cubes. These details are largely irrelevant to anyone who uses the MAPS and model LOGCUBE files, except to emphasize that users must be careful when interpreting the Voronoi-binned maps and spectra in these files. We provide guidance in this regard specific to the MAPS and model LOGCUBE files in Sections 12.1 and 12.2, respectively.

The DAP produces a MAPS and a model LOGCUBE file for each analysis approach, or DAPTYPE (Section 4), meaning there are two MAPS and model LOGCUBE files for each data cube successfully analyzed for DR15 (see Section 11.1). The detailed data models for the DAP output files are provided in Appendix C and via the DR15 website.[76]

Once the DAP has been executed on the individual data cubes, a final post-processing step is executed to construct a summary catalog called the DAPall catalog (Section 12.3). Currently, the primary intent of this catalog is to aid sample selection. We continue to improve the quantities provided by the DAPall catalog, but we currently do not recommend these data for scientific use without a detailed understanding of the data quality. In particular, note that the DAPall catalog does not provide *any* measurement uncertainties and only very simple methods are used to perform each measurement (e.g., the star formation rate does not account for attenuation). The DAPall summary catalog can be queried using both CASJobs[77] and Marvin. The full list of columns provided in the DAPall catalog is provided in Appendix C and via the DR15 website.[78]

12.1. MAPS Files

The MAPS file is the primary output file that provides the spaxel-by-spaxel quantities derived by the DAP. The measurements are organized in a series of extensions (Table 6) that contain images, or maps, with a format identical to the spatial dimensions of the DRP data cube. Extensions may contain a single map, like the measured stellar velocity, or a series of maps organized in "channels," like the fluxes derived for each emission line. When an extension contains more than one map, the channels are identified in the header. For example, the header of the extension containing the emission-line fluxes contains the header keyword and value `C19 = 'Ha-6564'`, indicating that the nineteenth channel contains the flux of the Hα line. Most extensions with DAP measurements have companion extensions with the inverse variance of the measurements and a quality mask (see Table 6).

Beyond this basic description of the data format, there are a few critical components of the MAPS files that users should keep in mind:

1. *Quantities are provided that a user must correct using the provided corrections.* In particular, the stellar velocity dispersions are provided as measured by pPXF, including the offset in spectral resolution between MaNGA and the MILES-HC template library; see Section 7.1.5. Similarly, the emission-line velocity dispersions are provided as would be determined by fitting a Gaussian directly to the emission line and must be corrected for the instrumental resolution; see Section 9.5. Finally, the spectral indices are provided as measured directly from the spectra and must be corrected for the effects of the velocity dispersion on the measurement; see Section 10.1.
2. *Basic quality assessments of the data are provided via bitmasks and should be used.* The DAP performs a number of quality checks during the measurement process. The mask bits triggered by the DAP modules are consolidated and incorporated into the bitmasks provided with the MAPS files (see Table 8). Any non-zero value of the bitmask indicates that the measurement should be either treated with care or ignored. Both the DAP source code and Marvin provide convenience Python classes that facilitate the use of the bitmasks[79] to appropriately flag the DAP data.
3. *Results for a binned spectrum are repeated for each spaxel in the bin.* When using *any* data from the VOR10-GAU-MILESHC files or results from the first three modules (Figure 3) of the HYB10-GAU-MILESHC files, the results in every spaxel do not necessarily represent unique measurements. This is critical to consider when, for example, fitting the data with a model or binning the data as a function of radius. The primary use of the BINID extension in the MAPS file (Table 6) is to allow users to select the unique measurements made for each mapped quantity.
4. *Flux units are per spaxel.* The units can be converted to surface brightness by multiplying by the pixel scale (i.e., four spaxels per arcsec^2). Integrations of the flux over map apertures can be done by summing spaxel values; however, be aware of the previous point about measurements being repeated for binned spaxels. The spectral stacking procedure is a simple average of the spaxels in each bin (Section 6.4), meaning that the units are correct; however, one should avoid apertures that do not enclose the full bin.

[76] Specifically, see the description at https://www.sdss.org/dr15/manga/manga-data/data-model/ and the detailed data model at https://data.sdss.org/datamodel/files/MANGA_SPECTRO_ANALYSIS/. A brief introduction for how to read the latter can be found at https://data.sdss.org/datamodel.

[77] https://skyserver.sdss.org/casjobs/

[78] The DAPall catalog is included in the list of MaNGA Catalogs at https://www.sdss.org/dr15/manga/manga-data/catalogs/, with the detailed datamodel at https://data.sdss.org/datamodel/files/MANGA_SPECTRO_ANALYSIS/DRPVER/DAPVER/dapall.html.

[79] https://www.sdss.org/dr15/algorithms/bitmasks/

5. *Velocities are offset by the input bulk redshift.* As discussed in Section 7.1.4, the velocities reported in the MAPS files have been offset by the input bulk redshift, z_0. These redshifts are most often provided by the NSA, and the value used for the bulk redshift is saved in km s^{-1} (cz_0) in the header keyword SCINPVEL. Because these bulk redshifts are not directly determined from the MaNGA data, they may not accurately offset the velocity to 0 km s^{-1} at the galaxy center. One can recover the redshift measured for each spaxel or binned spectrum, $z_{\rm obs}$, using Equation (5) in Section 7.1.4.
6. *Some velocity dispersion measurements are below the MaNGA instrumental resolution.* As mentioned above, both the stellar and emission-line velocity dispersions must be corrected; the former is corrected for the intentional offset between the template resolution and the galaxy data (Section 7.4.3) and the other is for the instrumental resolution of the data. It is possible to find measurements that are smaller than the quadrature correction due to the error distribution in the determination of either. The reason why we have left it to the users to decide how to treat these measurements is because the treatment may depend on the science goals. We discuss these measurements, in particular, and provide advice for their treatment in Section 7.7.
7. *Modeled properties of each emission line in a spectrum are not necessarily independent.* As executed for DR15, the DAP ties all emission-line velocities and ties the fluxes and velocity dispersion for many of the doublets; see Section 9.1 and Table 3. However, as discussed in Section 9.4, the relevant properties and errors are provided for each line, even if those properties are not independent parameters in the fit. Unfortunately, there is no way to determine which parameters are tied based solely on the provided output files; users must consult Table 3.
8. *Emission-line fluxes are corrected for Galactic foreground extinction but not dust within a given galaxy.* See point (4) in the next section.
9. *The non-parametric emission-line fluxes are provided largely as a check on the Gaussian modeling results.* The MAPS files provide both an EMLINE_GFLUX and EMLINE_SFLUX extension with the results from, respectively, the Gaussian model-fit results (Section 9.4) and the zeroth-moment measurements (Section 9.7). In general, the more precise measurements from the Gaussian modeling should be used. The two measurements only significantly differ at low flux levels (see Belfiore et al. 2019, Figure 4) when the Gaussian fit can be driven by noise. The rule of thumb is then to use the Gaussian results when it is similar to the non-parametric result; otherwise, do not trust either.
10. *Errors are based on formal calculations.* All errors provided by the DAP are currently based on either a formal calculation (see e.g., the description in the last paragraph of Section 7.5.1) or simple error propagation from the DRP. In general, we find these errors to be within a factor of 2 of the true error as determined by both idealized parameter-recovery simulations and empirical measurements of the random error using repeat observations. This is true of the stellar kinematics (Section 7.5), the model-fit emission-line properties (Belfiore et al. 2019), and the spectral indices (Section 10). In the latter two cases, we have provided simple prescriptions to recalibrate the provided formal errors to match our simulation results and/or empirical estimates; however, the user must apply these calibrations themselves.
11. *The mapped properties are covariant.* As we discuss throughout our paper (Sections 2, 6.2, 7.5.1; Figures 7–10), the MaNGA data cubes exhibit significant spatial covariance given the subsampling of the MaNGA 2.″5-diameter fiber beam into 0.″5 × 0.″5 spaxels. This covariance, of course, propagates to the derived parameters; however, we have not provided covariance matrices for the MAPS data in DR15. Initial simulations suggest that, to first order, the *correlation* matrix of the flux in a representative wavelength channel (e.g., Figure 7) is the same as for the derived parameters. In particular, a first-order correlation matrix for the derived quantities can be constructed assuming $\ln \rho_{jk} = -D_{jk}^2/7.37$, where j and k are the indices of two spaxels separated by a distance of D_{jk} in spaxels. Although the significant covariance between spaxels complicates the analysis of the data, it also allows one to visually assess the S/N$_g$ level at which DAP results for individual spaxels may suffer from systematic error: any results provide by the DAP that do not smoothly vary between adjacent spaxels are driven by systematic error in the associated algorithm.
12. *Visual inspection of the data can be critical.* The MaNGA dataset is incredibly varied. One may find interesting outliers or trends when looking for them; however, possibly not for the expected reason. For example, in searching for data cubes with the strongest gradient in D4000, one is led to find observations with more than one target in the IFU FOV, not individual galaxies with the strongest stellar population gradients. Both in the sense of looking at the mapped properties and the fitted spectra, conclusions should always be evaluated in the context of the original source data. Marvin is particularly useful for quick visual assessments of the data.[80]

12.2. Model LOGCUBE Files

The primary purpose of the model LOGCUBE file is to allow users to check the results of the two full-spectrum-fitting modules against the data, particularly when the data in the MAPS file appear to be unreasonable. The file is made up primarily of extensions that contain data cubes in the same format as the DRP LOGCUBE file that the DAP has analyzed. The full list of extensions and their content is provided in Table 7.

Points (2), (3), and (4) from the previous section on the MAPS file also apply when using the model LOGCUBE files. For the description of the bitmasks used in the model LOGCUBE file, see Table 9. In addition:

1. *The flux array provided in the* DAP *model* LOGCUBE *files is different from the flux array in the* DRP *data cube.* Although this is apparent from the data model of each of the relevant files, the naming convention of the two files can lead to confusion. For clarity, the model LOGCUBE

[80] https://dr15.sdss.org/marvin/

file always provides the *binned* spectra, and the name of the model LOGCUBE file always includes the DAPTYPE.

2. *The best-fit stellar continuum used to determine the stellar kinematics is not provided directly and must be constructed*. The models provided in the MODEL extension are the result of the combined continuum+emission-line fits performed by the emission-line module. To construct the best-fitting spectrum from the stellar kinematics module, one has to remove the emission lines (in extension EMLINE) and the difference between the stellar continuum determined between the two full-spectrum-fitting modules (in extension EMLINE_BASE). That is, the stellar continuum is computed as MODEL - EMLINE - EMLINE_BASE.[81]
3. *The models provided for the hybrid-binning scheme should be compared to the* DRP *data cube*. The model LOGCUBE file always provides the *binned* spectra. However, in the hybrid-binning scheme, the models provided have been fit to the individual spaxels because they are the result of the emission-line module. Although there are "binned spectra" composed of single spaxels, in general this means that the MODEL extension of the model LOGCUBE file for the hybrid-binning scheme (DAPTYPE = HYB10-GAU-MILESHC) must be compared to the DRP data cube, not its own FLUX array. The same is *not* true for the DAPTYPE = VOR10-GAU-MILESHC files.
4. *The spectra include Galactic extinction*. Section 6.5 notes that once the spectra are binned, the Galactic extinction is removed from the data and all spectral modeling and measurements provided by the last four modules of the DAP (see Figure 3) use extinction-corrected spectra. However, to facilitate the comparison of the models with the DRP-produced data cubes (particularly given the previous point), the extinction curve is reapplied to the data before being written to the model LOGCUBE file; the exact reddening correction applied is provided in the REDCORR extension.

12.3. The DAPall Summary Catalog

Similar to the MaNGA DRPall file, we provide a summary DAPall catalog that collates global information pulled or derived from the primary output files of the DAP. The DAPall catalog contains one row per PLATEIFU and DAPTYPE combination. As discussed in Section 4, the DAP analyzed 4731 observations using two analysis approaches for DR15, meaning the DAPall file has 9462 rows. However, a small fraction of those analysis attempts failed (see Section 11.1)—the failures are indicated by the DAPDONE column in the DAPall catalog. The most basic selection of rows from the DAPall catalog would then select observations that were successfully analyzed (DAPDONE == 1) for a given analysis approach (e.g., DAPTYPE == 'HYB10-GAU-MILESHC'). For convenience when querying properties in both the DRPall and DAPall catalogs, we also provide the row index in the DRPall database, DRPALLINDX, matched to the same observation (PLATEIFU).

We emphasize again (see the beginning of this section) that the current DAPall catalog is primarily provided as a convenience to aid sample selection. Roughly half of the DAPall columns contain either metadata pulled from the MAPS file headers that are relevant to the methods used in the analysis or metadata repeated from the DRPall catalog and provided for convenience. The other half are derived directly from the MAPS data with the aim of providing relevant quantities for queries based on the spatial coverage, S/N, redshift, internal kinematics, and composition of each galaxy. The methods that we have used to construct these data are simple, sometimes at the expense of performing the nuanced analysis needed for direct scientific use. In particular, no uncertainties are currently calculated for the properties unique to the DAPall catalog. Of course, any sample selection based on these quantities should be tempered by an understanding of the sample biases that may result (Wake et al. 2017), as well as the limitations in the measurement construction and return to the source MAPS data as necessary for a more nuanced analysis. Here, we briefly highlight some of these quantities and the details of their calculation.

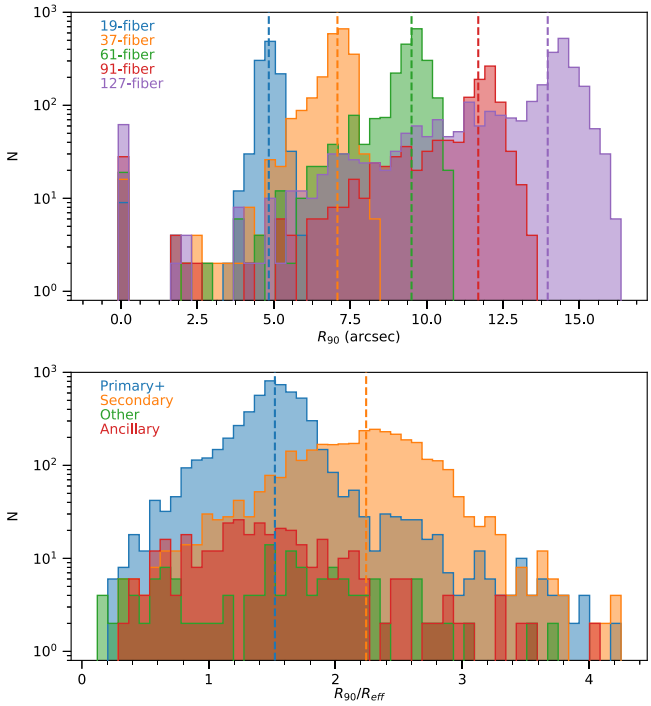

Figure 28. The radius to which at least 90% of a 2.″5 elliptical annulus is covered by spaxels analyzed by the DAP, R_{90}. The criteria selecting spaxels to be analyzed by the DAP is discussed in Section 6.1. The top panel shows the distribution of R_{90} in arcsec for observations taken with each bundle, colored by the bundle size. The bottom panel shows the distribution of R_{90} normalized by the elliptical-Petrosian half-light radius, R_{eff}, for galaxies belonging to the Primary+ and Secondary samples, as well as observations of ancillary or filler targets.

The MaNGA galaxy survey is designed with nominal radial coverage and S/N requirements (Yan et al. 2016a; Wake et al. 2017), and the DAPall file provides assessments of these quantities for each observation. For example, Figure 28 shows the distribution of the radial coverage for each observation in arcseconds and normalized by R_e. We define the radial coverage of each galaxy as the limiting radius to which at least 90% of the area of a 2.″5 elliptical annulus is observed by MaNGA spaxels. The distribution is as expected with median values that illustrate the on-sky size of the IFU and the designed 1.5 R_e and 2.5 R_e radial coverage of, respectively, the

[81] This convention is true for DR15, but the datamodel of the model LOGCUBE files will change in future releases.

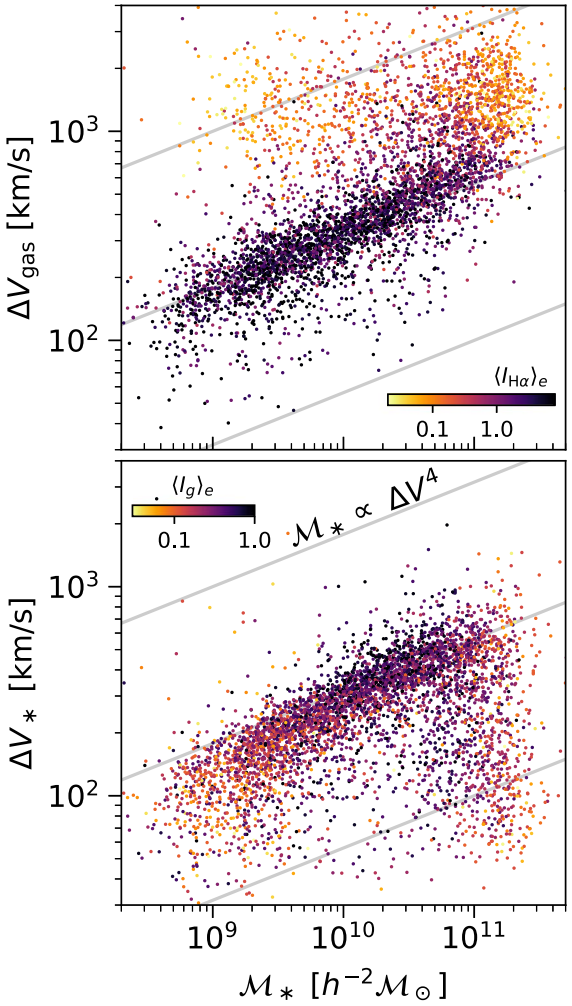

Figure 29. NSA stellar mass vs. the velocity gradient—defined as $\Delta V = (V_{\rm hi} - V_{\rm lo})/(1-(b/a)^2)^{1/2}$, where $V_{\rm hi}$ and $V_{\rm lo}$ are provided by the `DAPall` file—of the emission-line (top) and stellar (bottom) kinematics. Points are colored according to the mean surface brightness within 1 R_e.

Primary+ and Secondary samples of the main galaxy survey. The S/N metrics in the `DAPall` file provide, for example, the median S/N between 1 and 1.5R_e in the *griz* bands and the median *g*-band S/N for spectra between 0.0–1.0, 0.5–1.5, and 1.5–2.5R_e.

For global kinematic properties, the `DAPall` file provides simple measurements of the bulk redshift, velocity gradient, and velocity dispersion within 1R_e for the stellar and ionized gas tracers. Figure 29 shows our assessment of the velocity gradient for the gas and stars against the NSA stellar mass. This is a crude version of the Tully & Fisher (1977) (T-F) relation using ionized gas and stellar mass. We define the velocity gradient as the difference between the minimum and maximum measured velocity after removing 3σ outliers, and we apply a rough correction for the projection of the motions along the line of sight using the photometric ellipticity. That is, $\Delta V = (V_{\rm hi} - V_{\rm lo})(1-(b/a)^2)^{-1/2}$, where $V_{\rm hi}$ and $V_{\rm low}$ are provided in the `DAPall` catalog and b/a is provided by the `DRPall` catalog. Although this is a very basic assessment of the velocity field, we do find a correlation between the velocity gradient and stellar mass for galaxies with relatively high Hα and/or *g*-band surface brightness. The observed trend has roughly the expected form $\Delta V \propto M_*^4$ of the T-F relation.

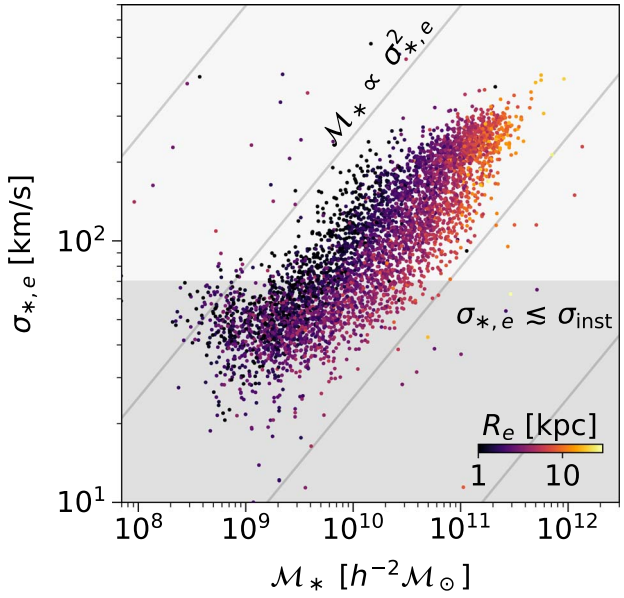

Figure 30. NSA stellar mass vs. light-weighted stellar velocity dispersion within 1 R_e from the `DAPall` file.

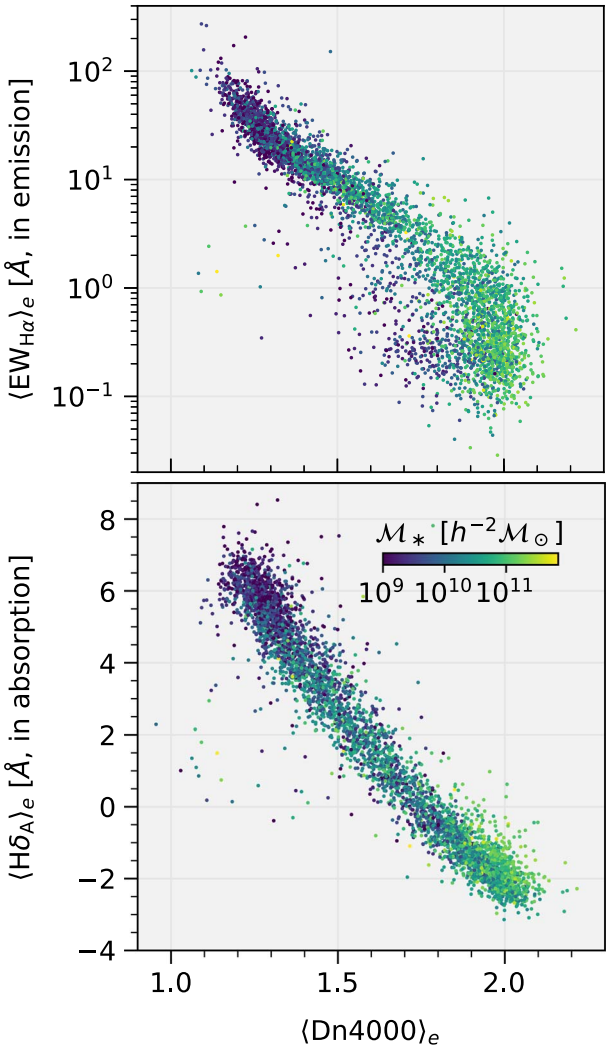

Figure 31. Dn4000 vs. the Hα equivalent width *in emission* (top) and the H$\delta_{\rm A}$ index (equivalent width in *absorption*) after subtracting the best-fitting emission-line model. Points are colored by NSA stellar mass.

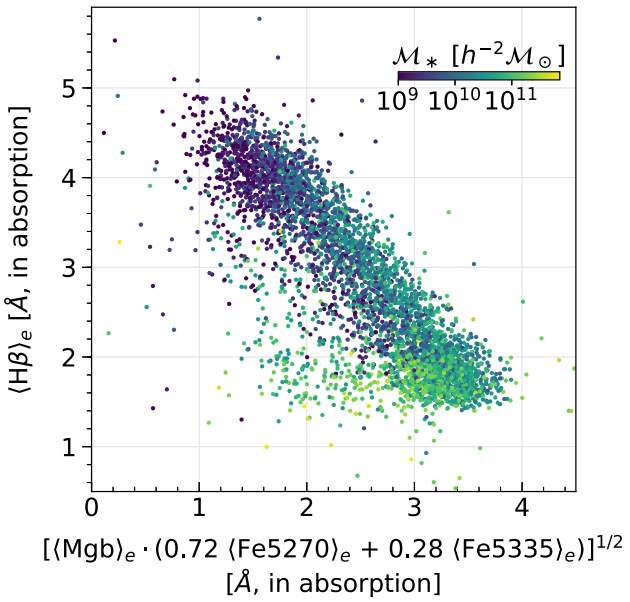

Figure 32. Typical stellar-population age–metallicity diagnostic using the Mgb, Fe5270, Fe5335, and Hβ absorption indices. Points are colored by NSA stellar mass.

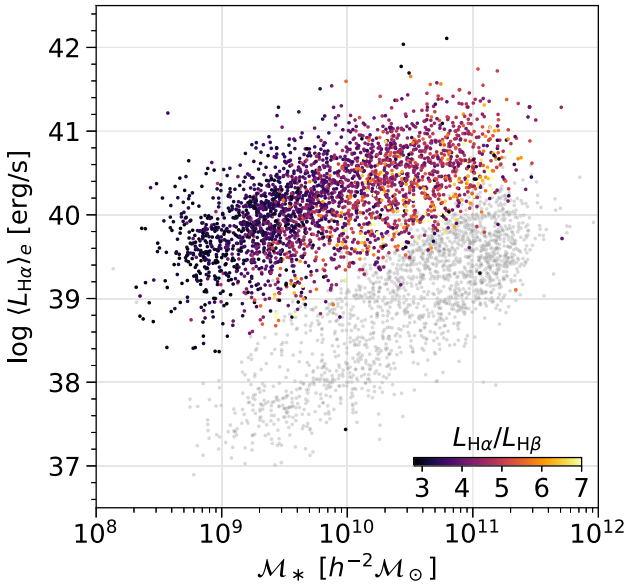

Figure 33. NSA stellar mass vs. the absolute luminosity in Hα. Points with Hα EW greater than 2 Å are colored by the Hα-to-Hβ luminosity ratio; others are set to gray.

Figure 30 shows the luminosity-weighted mean stellar velocity dispersion within $1R_e$ against the NSA stellar mass. In detail, we calculate $\sum_i I_{g,i}(\sigma^2_{\rm obs,}i - \delta\sigma^2_{{\rm inst},i})^{1/2}/\sum_i I_{g,i}$, where $I_{g,i}$ is the g-band-weighted mean flux for spaxel i, and the sum is over all i spaxels with luminosity-weighted bin (or individual spaxel) centers within $1R_e$. This is effectively the stellar-mass view of the Faber & Jackson (1976) relation, which is a projection of the ($\log M_*$, $\log \sigma_{*,e}$, $\log R_e$) mass plane (see Section 4 of Cappellari 2016, for a review) but it includes all morphological types instead of ETGs alone. As expected, the primary correlation between stellar and dynamical mass is evident. Above the instrumental resolution, the upper boundary envelope roughly follows the trend $M_* \propto \sigma^2_{*,e}$ observed for much smaller samples, with the expected flattening at larger masses ($M_* \sim 3 \times 10^{10}$) (see Figure 20 of Cappellari 2016). But the trend persists at lower σ_* (with a slight upturn at the lowest masses), confirming our ability to measure σ_* well below the MaNGA instrumental dispersion. A different projection of the MaNGA mass plane for the galaxies in the DR14 was presented in Li et al. (2018).

The `DAPall` file provides the unweighted median of all emission-line fluxes and equivalent widths, both from the Gaussian and non-parametric fits, and spectral indices for spectra with luminosity-weighted bin (or individual spaxel) centers within $1R_e$. These can be used as quick emission-line and stellar-population diagnostics, as demonstrated in Figures 31 and 32. Using the Hα flux measured within each IFU, we also provide a very rough estimate of the star formation rate: We calculate the absolute luminosity in Hα, using the luminosity distance—see the plot of $L_{H\alpha}$ against stellar mass in Figure 33—and adopt $\log {\rm SFR} = \log L_{H\alpha} - 41.27$ (Kroupa IMF; from the literature compilation provided by Kennicutt & Evans 2012).

13. Conclusions

We have presented a description of the MaNGA Data Analysis Pipeline (`DAP`), its output data products, and its performance. We recommend that readers who intend to use the provided data should closely read our "quick-start guide" in Section 2 and the description of the output products in Section 12 for particularly useful summary-level information and usage notes. In particular, Section 2 also serves as a guide to sections throughout this paper with detailed information regard each data product.

In its automated measurements of stellar kinematics, nebular emission-line properties, and spectral indices, the `DAP` is highly successful. In Section 7, we demonstrate via repeat observation and simulation that the `DAP` provides accurate stellar kinematics to ${\rm S/N}_g \sim 10$ to a minimum velocity dispersion of $\sim$50 km s^{-1}. The formally calculated errors in the stellar kinematics are very consistent with the direct estimates from repeat observations. In Belfiore et al. (2019) we show similar performance for our emission-line fitting module presented in Section 9. These two full-spectrum-fitting modules are shown to robustly fit the MaNGA spectra across its hugely varied dataset, owing much to the accuracy and fidelity of our data reduction and flux-calibration techniques (Law et al. 2016; Yan et al. 2016b). The few exceptions to this, as enumerated in Section 11.2.2, are the subject of ongoing improvements being made to the `DAP`.

A current drawback of the `DAP` products provided in DR15 is the limited wavelength range over which measurements are made (0.36–0.74 μm) owing to the wavelength coverage of the MILES stellar templates used in our full-spectrum fits. An exciting near-term development goal for the `DAP` will be the adoption of stellar templates from the MaStar Stellar Library (Yan et al. 2019), which samples a larger number of stars across a wider range of stellar parameters and over the full MaNGA wavelength range. While the MaStar library is not appropriate for stellar kinematics near the MaNGA instrumental resolution (since MaStar stars are *observed* with MaNGA itself), it will enable measurements of stellar- and gas-phase spectral features as red as 1.0 μm and provide the

basis for new stellar-population-synthesis models (C. Maraston et al. 2019, in preparation).

The DAP design philosophy has included a focus on measurements that can be made directly from individual MaNGA spectra. These measurements are generic to galaxy spectra at similar wavelengths and spectral resolution. Indeed, while the goals of an automated pipeline require fine-tuning to the MaNGA dataset and data format, the DAP sub-routines have been written with generality in mind to ensure that they can be adapted for other datasets.

Future versions or extensions to the DAP may incorporate estimates of higher-level "model-derived" quantities. Natural extensions include continuum fitting in order to derive stellar population properties (e.g., stellar age and metallicity) and multi-line analysis of gas-phase emission lines for estimates of ionization and gas-phase metallicity.

Future extensions in the context of MaNGA or other IFU data might also move beyond the independent treatment of spectra from each spatial bin and fit models that attempt to capture the spatial information in each galaxy's data cube. These could include dynamical models of DAP-derived kinematic maps, as well as forward models of spatially dependent stellar or gas-phase galaxy components. The DAP currently inherits structural information (e.g., the galaxy's effective radius) from extant photometric catalogs. Spatially dependent modeling might instead be iterative, making use of the structural information present in the DAP output maps themselves.

We sincerely appreciate the effort of our referee in providing a thoughtful, constructive, and thorough report that was of great benefit to our paper. M.C. acknowledges support from a Royal Society University Research Fellowship. M.A.B. acknowledges NSF Award AST-1517006. C.A.T. acknowledges NSF Award AST-1554877. Z.Z. is supported by the National Natural Science Foundation of China No. 11703036. M.Y. gratefully acknowledges the financial support from China Scholarship Council (CSC). This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013); numpy (Oliphant 2007); scipy (Jones et al. 2001); and matplotlib (Hunter 2007).

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS website is available at www.sdss.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

Appendix A
Spectral-resolution Matching

To match the resolution of a template library to that of the DRP-produced spectra, we convolve the discretely sampled flux density, $f(\lambda)$ in units of $10^{-17}\,\mathrm{erg\,s^{-1}\,cm^{-2}\,\AA^{-1}}$, with a Gaussian kernel, $g(\lambda)$, where the standard deviation of the kernel is a function of the wavelength, $\sigma_\lambda(\lambda)$. In general, the convolution is defined as:

$$(f^{*}g)(\lambda) = \int_{-\infty}^{\infty} f(\Lambda)\; g(\lambda-\Lambda,\,\sigma_g)\; d\Lambda = \int_{-\infty}^{\infty} \frac{f(\Lambda)}{\sqrt{2\pi}\;\sigma_g(\Lambda)} \exp\left(-\frac{(\lambda-\Lambda)^2}{2\;\sigma_g(\Lambda)^2}\right) d\Lambda. \quad (30)$$

It is important to note that the integral of the kernel is normalized to unity. In practice, the application of Equation (30) should account for the discrete sampling and censoring of the data over the observed spectral range. We do so by normalizing the convolution by the integral over the kernel, which is significantly different from unity only near the edges of the observed spectral range. We additionally simplify Equation (30) by performing the convolution in pixel space, converting σ_λ to σ_p. Therefore, the convolved spectrum at wavelength λ_i becomes the kernel-weighted mean of the spectrum over all pixels, with the kernel centered at Λ_j:

$$(f^{*}g)(\lambda_i) = \frac{\sum_j f(\Lambda_j) g(\lambda_i-\Lambda_j,\,\sigma_j)}{\sum_j g(\lambda_i-\Lambda_j,\,\sigma_j)}. \quad (31)$$

By adopting a Gaussian LSF for both the spectral templates and the MaNGA data, we determine the kernel parameters, $\sigma_p(\lambda)$, as follows. We define the spectral resolution as $R = \lambda/\Delta\lambda$, where $\Delta\lambda$ is the FWHM of the spectral resolution element. The standard deviation of the resolution element in angstroms is then $\sigma_\lambda = \frac{\lambda}{n_\sigma R}$, where $n_\sigma = \Delta\lambda/\sigma_\lambda \sim 2.35$ for a Gaussian LSF.

For two spectra with spectral resolutions $R_1 \geqslant R_2$, the defining parameters of the Gaussian LSFs can be related by

$$\sigma^2_{\lambda,2} = \sigma^2_{\lambda,1} + \sigma^2_{\lambda,d}, \quad (32)$$

where we define

$$\sigma^2_{\lambda,d} \equiv \left(\frac{\lambda}{f}\right)^2 (R_2^{-2} - R_1^{-2}). \quad (33)$$

For the application of Equation (31), we convert the units of $\sigma^2_{\lambda,d}$ from wavelength to pixels. Performing the convolution in pixel units has the added advantage that it allows for similar application of Equation (31) to spectra that are either sampled linearly or geometrically—sampled in linear steps of $\log_b \lambda$—in

wavelength. If linearly sampled,

$$\sigma^2_{p,d} \equiv \left(\frac{\lambda}{f\ \delta\lambda}\right)^2 (R_2^{-2} - R_1^{-2}) \tag{34}$$

where $\delta\lambda$ is the pixel scale in angstroms. If geometrically sampled, the pixel size is converted to velocity,

$$\delta v = c\ \ln b\ (\delta \log_b \lambda), \tag{35}$$

such that

$$\sigma^2_{p,d} = \sigma^2_{v,d}(\delta v)^{-2} \tag{36}$$

$$= \left(\frac{c}{\lambda}\right)^2 \sigma^2_{\lambda,d}(\delta v)^{-2}$$

$$= \left(\frac{c}{f\ \delta v}\right)^2 (R_2^{-2} - R_1^{-2}), \tag{37}$$

where c is the speed of light in km s^{-1}.

As stated earlier, Equation (32) assumes $R_1 \geqslant R_2$. However, in practice, some spectral libraries may not have resolutions that are larger than the MaNGA data over the full spectral range. In our resolution matching algorithm, we define a minimum value of $\sigma_{p,d}$, ϵ_σ, below which we approximate the Gaussian kernel as a Kronecker delta function. Therefore, as long as

$$\sigma_{p,d} \equiv \sigma^2_{p,d} \Big/ \sqrt{|\sigma^2_{p,d}|} \geqslant -\epsilon_\sigma, \tag{38}$$

the behavior of the convolution should not be affected.

However, we typically set $\epsilon_\sigma = 0$ and we need to robustly handle regions where $R_1 < R_2$. For these spectral regions, we highlight three approaches:

1. Trim the spectral range to only those spectral regions where the existing resolution is better than the target resolution,
2. Match the existing resolution to the target resolution up to some constant offset that must be accounted for in subsequent analyses, or
3. Allow for a wavelength dependent difference in the spectral resolution that must be accounted for in subsequent analyses.

Our code allows for selection of the first or second approach. Our standard practice is currently to adopt the first approach; our code does not allow for the third option.

In the first approach, pixels with $\sigma_{p,d} < -\epsilon_\sigma$ are masked from subsequent analysis, and the convolution algorithm does not alter the spectral resolution of these pixels.

In the second approach, we define

$$\sigma^2_{v,o} = -\min(\sigma^2_{v,d}) - \max(\epsilon_\sigma \delta v)^2 \tag{39}$$

where δv is constant for the geometrically binned spectrum and is wavelength dependent, $\delta v = c\delta\lambda/\lambda$, for the linearly binned spectra. If $\sigma^2_{v,o} > 0.0$, it must be that $\min(\sigma^2_{v,d}) < -\max(\epsilon_\sigma \delta v)^2$ such that an offset should be applied. In that case, the returned kernel parameters are

$$\sigma'_{v,d} = \sqrt{\sigma^2_{v,d} + \sigma^2_{v,o}} \tag{40}$$

with the units converted to pixels using Equation (36). In this approach, no pixels are masked and $\sqrt{\sigma^2_{v,o}}$ is returned for use in subsequent analysis. Otherwise (i.e., Equation (39) yields $\sigma^2_{v,o} \leqslant 0.0$), the returned offset is set to zero.

It should be noted that the offset, $\sigma_{v,o}$, is always kept in units of km s^{-1}, regardless of the spectral sampling. This facilitates later adjustment of the offset by a constant Gaussian velocity dispersion. This is useful for imposing a single offset for spectral templates at different resolutions (impose the maximum $\sigma^2_{v,o}$ on all spectral templates), or to apply a constant velocity dispersion offset for kinematic reasons.

Appendix B
Propagation of Spectral-resolution Errors in the σ_* Error Budget

Assuming that all instrumental LSFs and the stellar LOSVD are Gaussian, we can write the pPXF-measured stellar velocity dispersion as (see Equation (6)):

$$\sigma^2_{\rm obs} = \sigma^2_* + \sigma^2_{\rm g} - \sigma^2_{\rm t} \tag{41}$$

where $\delta\sigma_{\rm inst} = \sigma^2_{\rm g} - \sigma^2_{\rm t}$ is the quadrature difference in the instrumental resolution of the galaxy spectra, $\sigma_{\rm g}$, and the template spectra, $\sigma_{\rm t}$. Here, we explore how the errors in $\sigma_{\rm g}$ and $\sigma_{\rm t}$ propagate to the error in $\sigma_{\rm obs}$ at fixed σ_*, with the aim of minimizing the influence of instrumental-dispersion uncertainties.

By assuming that all errors are Gaussian and that one can accurately perform nominal error propagation, we can write

$$\left(\frac{\epsilon[\delta\sigma^2_{\rm inst}]}{\delta\sigma^2_{\rm inst}}\right)^2 = \left(\frac{\epsilon[\sigma^2_{\rm g}]}{\sigma^2_{\rm g} - \sigma^2_{\rm t}}\right)^2 + \left(\frac{\epsilon[\sigma^2_{\rm t}]}{\sigma^2_{\rm g} - \sigma^2_{\rm t}}\right)^2, \tag{42}$$

where $\epsilon[x]$ is the formal error in x. Adopting $\xi = \sigma_{\rm g}/\sigma_{\rm t}$ and

$$\frac{\epsilon[\sigma_{\rm inst}]}{\sigma_{\rm inst}} \equiv \frac{\epsilon[\sigma_{\rm g}]}{\sigma_{\rm g}} \sim \frac{\epsilon[\sigma_{\rm t}]}{\sigma_{\rm t}}, \tag{43}$$

we can simplify Equation (42) to

$$\frac{\epsilon[\delta\sigma_{\rm inst}]}{\delta\sigma_{\rm inst}} = \frac{\epsilon[\sigma_{\rm inst}]}{\sigma_{\rm inst}} \frac{\sqrt{\xi^4 + 1}}{\xi^2 - 1}, \tag{44}$$

where $\xi > 1$. From Equation (6), we can derive the error in $\sigma_{\rm obs}$ at fixed σ_*,

$$\frac{\epsilon[\sigma_{\rm obs}]}{\sigma_{\rm obs}} = \frac{\delta\sigma^2_{\rm inst}}{\sigma^2_{\rm obs}} \frac{\epsilon[\delta\sigma_{\rm inst}]}{\delta\sigma_{\rm inst}}. \tag{45}$$

Finally, substituting Equation (44) and rewriting in terms of σ_*, $\sigma_{\rm g}$, and ξ, we find:

$$\frac{\epsilon[\sigma_{\rm obs}]}{\sigma_{\rm obs}} = \frac{\epsilon[\sigma_{\rm inst}]}{\sigma_{\rm inst}} \frac{\sqrt{\xi^4 + 1}}{\xi^2(\sigma^2_*/\sigma^2_{\rm g} + 1) - 1}. \tag{46}$$

Defining

$$\Phi \equiv \frac{\sigma_{\rm inst}}{\sigma_{\rm obs}} \frac{\epsilon[\sigma_{\rm obs}]}{\epsilon[\sigma_{\rm inst}]} = \frac{\sqrt{\xi^4 + 1}}{\xi^2(\sigma^2_*/\sigma^2_{\rm g} + 1) - 1}, \tag{47}$$

we show Φ as a function of $\xi \equiv \sigma_{\rm g}/\sigma_{\rm t}$ for five discrete ratios of the target stellar velocity dispersion to the instrumental

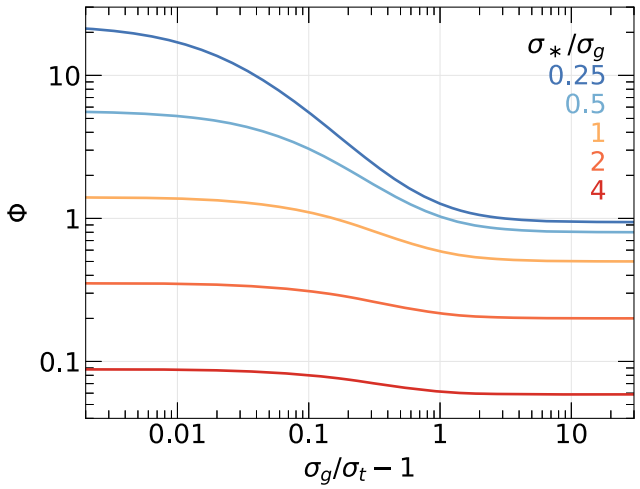

Figure 34. Ratio of the fractional error in σ_{obs} to the fractional error in σ_{inst} (i.e., Φ; Equation (47)) at fixed σ_* as a function of the ratio between the instrumental resolution of the galaxy spectra σ_g and the template spectra σ_t. Note that one *always* improves the fractional error in σ_{obs} by using higher resolution templates to measure the stellar kinematics.

resolution of the galaxy observations in Figure 34. In accordance with intuition, Figure 34 shows that one can limit the influence of σ_{inst} errors on σ_{obs} by performing observations at higher spectral resolution; i.e., Φ decreases as σ_*/σ_g increases. Interestingly, Figure 34 also shows that one can *always* reduce the influence of instrumental resolution errors by fitting the galaxy spectra with intrinsically higher resolution templates. Moreover, it shows that the relevant reduction of Φ increases as σ_*/σ_g decreases; i.e., as one attempts to measure σ_* toward and below the instrumental resolution of the galaxy data.

The limiting values of Φ when $\xi = 1$ and as $\xi \to \infty$ are, respectively,

$$\Phi_{\xi=1} = \frac{\sqrt{2}}{\sigma_*^2/\sigma_g^2} \tag{48}$$

and

$$\Phi_{\xi\to\infty} = \frac{1}{\sigma_*^2/\sigma_g^2 + 1}, \tag{49}$$

such that

$$\frac{\Phi_{\xi=1}}{\Phi_{\xi\to\infty}} = \sqrt{2} + \frac{1}{\sigma_*^2/\sigma_g^2}. \tag{50}$$

Therefore, improvements in Φ can be more than a factor of 2 when $\sigma_* \sim \sigma_g$ simply by using higher resolution templates.

Finally, we note from Figure 34 that ξ need not be too large to make significant gains. Even with the modest 16% difference in resolution between MaNGA and MILES, Equation (46) yields $\Phi_{\xi=1}/\Phi = 1.43$ when $\sigma_* = \sigma_g$. In fact, the asymptotic behavior of Φ is such that one expects diminishing returns for $\xi \equiv \sigma_g/\sigma_t$ larger than a factor of ~2–3.

Appendix C
MaNGA DAP Data Model

This appendix provides the DAP data model via a series of tables:

1. The two primary output files of the DAP (Section 12) are the MAPS and model LOGCUBE files. The names and content of each extension in these files are provided in Tables 6 and 7, respectively.
2. Both files include bitmask extensions.[82] The relevant bitmask types for the MAPS and model LOGCUBE files are MANGA_DAPPIXMASK and MANGA_DAPSPECMASK, respectively. The bit values, names, and descriptions are provided in Tables 8 and 9. The global quality assessment of a DAP file is provided by the MANGA_DAPQUAL bitmask type (given as the DAPQUAL header keyword in the primary extension of each file) with bit values, names, and descriptions provided in Table 10.
3. Finally, the DAPall file (Section 12.3) is a summary catalog of the DAP parameters and global quantities based on the output data. The tabulated data that it provides is listed in Table 11.

[82] For an introduction to SDSS bitmasks and further documentation of the DAP mask bits, see https://www.sdss.org/dr15/algorithms/bitmasks/.

Table 6
DAP MAPS File Extensions

Index	Name	Channels	Units	Description
0	PRIMARY	0	⋯	Empty extension with primary header information.
		Coordinates and Binning		
1	SPX_SKYCOO	2	arcsec	Sky-right offsets—$+x$ toward +R.A. and $+y$ toward +decl.—of each spaxel from the galaxy center.
2	SPX_ELLCOO	3	arcsec, unitless, deg	Elliptical polar coordinates of each spaxel from the galaxy center: R, R/R_e, θ. In the limit of a tilted thin disk, these are the in-plane disk radius and azimuth; the second channel is the radius normalized by the elliptical-Petrosian effective radius from the NSA.
3	SPX_MFLUX	1	10^{-17} erg s^{-1} cm^{-2} Å^{-1} spaxel^{-1}	The g-band-weighted mean flux, *not* corrected for Galactic extinction or internal attenuation.
4	SPX_MFLUX_IVAR	1		Inverse variance of g-band-weighted mean flux.
5	SPX_SNR	1		Mean g-band weighted signal-to-noise ratio per pixel.
6	BINID	5		Numerical ID for spatial bins for the binned spectra, stellar-continuum results, emission-line moment results, emission-line model results, and spectral-index results.
7	BIN_LWSKYCOO	2	arcsec	Light-weighted sky-right offsets—$+x$ toward +R.A. and $+y$ toward +decl.—of each bin from the galaxy center.
8	BIN_LWELLCOO	3	arcsec, unitless, deg	Light-weighted elliptical polar coordinates of each bin from the galaxy center: R, R/R_e, θ. In the limit of a tilted thin disk, these are the in-plane disk radius and azimuth; the second channel is the radius normalized by the elliptical-Petrosian effective radius from the NSA.
9	BIN_AREA	1	arcsec^2	Area of each bin.
10	BIN_FAREA	1		Fractional area that the bin covers for the expected bin shape (only relevant for radial binning).
11	BIN_MFLUX	1	10^{-17} erg s^{-1} cm^{-2} Å^{-1} spaxel^{-1}	The g-band-weighted mean flux for the binned spectra, *not* corrected for Galactic extinction or internal attenuation.
12	BIN_MFLUX_IVAR	1		Inverse variance of g-band-weighted mean flux for the binned spectra.
13	BIN_MFLUX_MASK	1		Bit mask for the g-band-weighted mean flux per bin.
14	BIN_SNR	1		Mean g-band-weighted signal-to-noise ratio per pixel in the binned spectra.
		Stellar Kinematics		
15	STELLAR_VEL	1	km s^{-1}	Line-of-sight stellar velocity, relative to the input guess redshift (given as cz by the SCINPVEL keyword in the header of the PRIMARY extension and most often identical to the NSA redshift).
16	STELLAR_VEL_IVAR	1		Inverse variance of stellar velocity measurements.
17	STELLAR_VEL_MASK	1		Data quality mask for stellar velocity measurements.
18	STELLAR_SIGMA	1	km s^{-1}	Raw line-of-sight stellar velocity dispersion (must be corrected using STELLAR_SIGMACORR to obtain the astrophysical dispersion).
19	STELLAR_SIGMA_IVAR	1		Inverse variance of stellar velocity dispersion.
20	STELLAR_SIGMA_MASK	1		Data quality mask for stellar velocity dispersion.
21	STELLAR_SIGMACORR	1	km s^{-1}	Quadrature correction for STELLAR_SIGMA to obtain the astrophysical velocity dispersion.
22	STELLAR_CONT_FRESID	2		68% and 99% growth of the fractional residuals between the model and data.
23	STELLAR_CONT_RCHI2	1		Reduced chi-square of the stellar continuum fit.
		Emission-line Properties		
24	EMLINE_SFLUX	22	10^{-17} erg s^{-1} cm^{-2}/spaxel^{-1}	Non-parametric summed flux *after subtracting the stellar-continuum model*. The emission-line fluxes account for Galactic reddening using the $E(B-V)$ value (copied to the DAP primary headers, see EBVGAL) provided by the DRP header and assuming the reddening law provided by O'Donnell (1994); however, no attenuation correction is applied due to dust internal to the galaxy.
25	EMLINE_SFLUX_IVAR	22		Inverse variance for summed flux measurements.
26	EMLINE_SFLUX_MASK	22		Data quality mask for summed flux measurements.
27	EMLINE_SEW	22	Å	Non-parametric equivalent widths measurements (based one EMLINE_SFLUX).
28	EMLINE_SEW_IVAR	22		Inverse variance for non-parametric equivalent-width measurements.
29	EMLINE_SEW_MASK	22		Data quality mask for non-parametric equivalent-width measurements.
30	EMLINE_GFLUX	22	10^{-17} erg s^{-1} cm^{-2}/spaxel^{-1}	

Table 6
(Continued)

Index	Name	Channels	Units	Description
				Gaussian profile integrated flux *from a combined continuum+emission-line fit. The flux ratio of the [O III], [O I], and [N II] lines are fixed and cannot be treated as independent measurements.* The emission-line fluxes account for Galactic reddening using the $E(B-V)$ (copied to the DAP primary headers, see EBVGAL) value provided by the DRP header and assuming the reddening law provided by O'Donnell (1994); however, no attenuation correction is applied due to dust internal to the galaxy.
31	EMLINE_GFLUX_IVAR	22		Inverse variance for Gaussian flux measurements.
32	EMLINE_GFLUX_MASK	22		Data quality mask for Gaussian flux measurements.
33	EMLINE_GEW	22	Å	Gaussian-fitted equivalent-width measurements (based on EMLINE_GFLUX).
34	EMLINE_GEW_IVAR	22		Inverse variance for Gaussian-fitted equivalent-width measurements.
35	EMLINE_GEW_MASK	22		Data quality mask for Gaussian-fitted equivalent-width measurements.
36	EMLINE_GVEL	22	km s^{-1}	Line-of-sight emission-line velocity, relative to the input guess redshift (given as *cz* by the SCINPVEL keyword in the header of the PRIMARY extension and most often identical to the NSA redshift). A velocity is provided for each line, *but the velocities are identical for all lines* because the parameters are tied during the fitting process.
37	EMLINE_GVEL_IVAR	22		Inverse variance for Gaussian-fitted velocity measurements, which are *the same for all lines and should not be combined as if independent measurements*.
38	EMLINE_GVEL_MASK	22		Data quality mask for Gaussian-fitted velocity measurements.
39	EMLINE_GSIGMA	22	km s^{-1}	Gaussian profile velocity dispersion as would be measured from a direct Gaussian fit (must be corrected using EMLINE_INSTSIGMA to obtain the astrophysical dispersion). *The velocity dispersions of the [O II], [O III], [O I], and [N II] lines are tied and cannot be treated as independent measurements.*
40	EMLINE_GSIGMA_IVAR	22		Inverse variance for Gaussian profile velocity dispersion.
41	EMLINE_GSIGMA_MASK	22		Data quality mask for Gaussian profile velocity dispersion.
42	EMLINE_INSTSIGMA	22	km s^{-1}	The instrumental dispersion at the fitted center of each emission line.
43	EMLINE_TPLSIGMA	22	km s^{-1}	The dispersion of each emission line used in the template spectra.
		Spectral Indices		
44	SPECINDEX	46	Å, mag	Spectral-index measurements.
45	SPECINDEX_IVAR	46		Inverse variance for spectral index maps.
46	SPECINDEX_MASK	46		Data quality mask for spectral index maps.
47	SPECINDEX_CORR	46	unitless, mag	Corrections to apply to account for the velocity dispersion and effectively determine the index without Doppler broadening.

Table 7
DAP Model LOGCUBE File Extensions

Index	Name	Units	Description
0	PRIMARY	⋯	Empty extension with primary header information.
1	FLUX	10^{-17} erg s^{-1} cm^{-2} Å^{-1} spaxel^{-1}	Flux of the *binned* spectra.
2	IVAR	⋯	Inverse variance in the binned spectra.
3	MASK	⋯	Bitmask for the binned and model spectra.
4	WAVE	Å	Vacuum-wavelength vector.
5	REDCORR	⋯	Reddening correction applied during the fitting procedures; calculate the de-reddened flux as FLUX * REDCORR.
6	MODEL	10^{-17} erg s^{-1} cm^{-2} Å^{-1} spaxel^{-1}	The best fitting model spectra (sum of the fitted continuum and emission-line models).
7	EMLINE	10^{-17} erg s^{-1} cm^{-2} Å^{-1} spaxel^{-1}	The model spectrum with *only* the emission lines.
8	EMLINE_BASE	10^{-17} erg s^{-1} cm^{-2} Å^{-1} spaxel^{-1}	The adjustment to the stellar continuum made during the combined continuum+emission-line fit.
9	EMLINE_MASK	⋯	The bitmask that only applies to the emission-line modeling.
10	BINID	⋯	Numerical ID for spatial bins in 5 channels: (1) binned spectra, (2) stellar-continuum results, (3) empty, (4) emission-line model results, and (5) empty; i.e., channels 1, 2, and 4 are the same as the BINID extension in the MAPS files and channels 3 and 5 are empty.

Table 8
DAP Mapped Quantity Mask Bits (MANGA_DAPPIXMASK)

Bit ($\log_2$)	Name	Description
0	NOCOV	No coverage in this spaxel.
1	LOWCOV	Low coverage in this spaxel.
2	DEADFIBER	Major contributing fiber is dead.
3	FORESTAR	A foreground star influences the flux in this spaxel.
4	NOVALUE	Spaxel ignored by the DAP.
5	UNRELIABLE	Value is deemed unreliable.
6	MATHERROR	A mathematical error occurred when computing the value.
7	FITFAILED	Fit to this spaxel failed.
8	NEARBOUND	Fitted value is too near an imposed boundary.
9	NOCORRECTION	Appropriate correction is not available.
10	MULTICOMP	A multi-component velocity feature has been detected.
30	DONOTUSE	Do not use this spaxel for science.[a]

Note.
[a] This bit is a consolidation of all spaxels flagged as NOCOV, LOWCOV, DEADFIBER, FORESTAR, NOVALUE, MATHERROR, FITFAILED, or NEARBOUND.

Table 9
DAP Model LOGCUBE Mask Bits (MANGA_DAPSPECMASK)

Bit ($\log_2$)	Name	Description
0	IGNORED	Pixel ignored.
1	FORESTAR	A foreground star influences the flux in this spaxel.
2	FLUXINVALID	Invalid flux measurements in pixel.
3	IVARINVALID	Invalid inverse variance in pixel ($\leqslant$0).
4	ARTIFACT	Flux measurements affected by a designated artifact.
5	FITIGNORED	Pixel not included in the relevant fit.
6	FITFAILED	Fit to spectral region failed.
7	ELIGNORED	Pixel ignored during emission-line fit.
8	ELFAILED	Fit to emission-line in this spectral region failed.

Table 10
DAP Quality Mask Bits (MANGA_DAPQUAL)

Bit ($\log_2$)	Name	Description
0	FORESTAR	A foreground star is present within the data cube field-of-view.
1	BADZ	Mismatch between redshifts derived from MaNGA observations and provided by the NASA-Sloan Atlas.
2	LINELESS	No significant nebular emission detected.
3	PPXFFAIL	pPXF fails to fit this object.
4	SINGLEBIN	Voronoi-binning algorithm forced all spectra into a single bin.
5	BADGEOM	Invalid input geometry; elliptical coordinates and effective radius are meaningless.
28	DRPCRIT	Critical failure in DRP.
29	DAPCRIT	Critical failure in DAP.
30	CRITICAL	Critical failure in DRP or DAP.

Table 11
DAPall Table Data

Column	Units	Description
PLATE	⋯	Plate number.
IFUDESIGN	⋯	IFU design number.
PLATEIFU	⋯	String combination of PLATE-IFU to ease searching.
MANGAID	⋯	MaNGA ID string.
DRPALLINDX	⋯	Row index of the observation in the DRPallfile.
MODE	⋯	3D mode of the DRPfile (CUBE or RSS).
DAPTYPE	⋯	Keyword of the analysis approach used (e.g., HYB10-GAU-MILESHC).
DAPDONE	⋯	Flag that MAPS file successfully produced.
OBJRA	deg	R.A. of the galaxy center.
OBJDEC	deg	Decl. of the galaxy center.
IFURA	deg	R.A. of the IFU pointing center (generally the same as OBJRA).
IFUDEC	deg	Decl. of the IFU pointing center (generally the same as OBJDEC).
MNGTARG1	⋯	Main survey targeting bit.
MNGTARG2	⋯	Non-galaxy targeting bit.
MNGTARG3	⋯	Ancillary targeting bit.
Z	⋯	Redshift used to set initial guess velocity (typically identical to NSA_Z).
LDIST_Z	h^{-1} Mpc	Luminosity distance, D_L, based on Z and a standard cosmology.[a]
ADIST_Z	h^{-1} Mpc	Angular-diameter distance, D_A, based on Z and a standard cosmology.[a]
NSA_Z	⋯	Redshift from the NASA-Sloan Atlas (NSA).
NSA_ZDIST	⋯	NSA distance estimate using peculiar velocity model of Willick et al. (1997); multiply by c/H_0 for Mpc.
LDIST_NSA_Z	h^{-1} Mpc	Luminosity distance based on NSA_Z and a standard cosmology.[a]
ADIST_NSA_Z	h^{-1} Mpc	Angular-diameter distance based on NSA_Z and a standard cosmology.[a]
NSA_ELPETRO_BA	⋯	NSA isophotal axial ratio from an elliptical-Petrosian analysis of the *r*-band image.
NSA_ELPETRO_PHI	deg	NSA isophotal position angle from an elliptical-Petrosian analysis of the *r*-band image.
NSA_ELPE-TRO_TH50_R	arcsec	Half-light radius provided by the NSA from an elliptical-Petrosian analysis of the *r*-band image; this is the same as R_e below.
NSA_SERSIC_BA	⋯	NSA isophotal axial ratio from Sérsic fit to the *r*-band image.
NSA_SERSIC_PHI	deg	NSA isophotal position angle from Sérsic fit to the *r*-band image.
NSA_SERSIC_TH50	arcsec	NSA effective radius from the Sérsic fit to the *r*-band image.
NSA_SERSIC_N	⋯	NSA Sérsic index from the Sérsic fit to the *r*-band image.
VERSDRP2	⋯	Version of DRP used for 2d reductions.
VERSDRP3	⋯	Version of DRP used for 3d reductions.
VERSCORE	⋯	Version of mangacore used by the DAP.
VERSUTIL	⋯	Version of idlutils used by the DAP.
VERSDAP	⋯	Version of mangadap.
DRP3QUAL	⋯	DRP 3D quality bit.
DAPQUAL	⋯	DAP quality bit.
RDXQAKEY	⋯	Configuration keyword for the method used to assess the reduced data.
BINKEY	⋯	Configuration keyword for the spatial-binning method.
SCKEY	⋯	Configuration keyword for the method used to model the stellar-continuum.
ELMKEY	⋯	Configuration keyword that defines the emission-line moment measurement method.
ELFKEY	⋯	Configuration keyword that defines the emission-line modeling method.
SIKEY	⋯	Configuration keyword that defines the spectral-index measurement method.
BINTYPE	⋯	Type of binning used.
BINSNR	⋯	Target for bin S/N, if Voronoi binning.
TPLKEY	⋯	The identifier of the template library, e.g., MILES.
DATEDAP	⋯	Date the DAP file was created and/or last modified.
DAPBINS	⋯	The number of "binned" spectra analyzed by the DAP.
RCOV90	arcsec	

Table 11
(Continued)

Column	Units	Description
		Semimajor-axis radius (R) below which spaxels cover at least 90% of elliptical annuli with width $R \pm 2\farcs5$. This should be independent of the DAPTYPE.
SNR_MED	...	Median S/N per pixel in the "*griz*" bands within 1.0–1.5 R_e. This should be independent of the DAPTYPE.
SNR_RING	...	S/N in the "*griz*" bands when binning all spaxels within 1.0–1.5 R_e. This should be independent of the DAPTYPE.
SB_1RE	10^{-17} erg s^{-1} cm^{-2} Å^{-1} spaxel^{-1}	Mean g-band surface brightness of valid spaxels within 1 R_e. This should be independent of the DAPTYPE.
BIN_RMAX	R_e	Maximum g-band luminosity-weighted semimajor radius of any "valid" binned spectrum.
BIN_R_N	...	Number of binned spectra with g-band luminosity-weighted centers within 0–1, 0.5–1.5, and 1.5–2.5 R_e.
BIN_R_SNR	...	Median g-band S/N of all binned spectra with luminosity-weighted centers within 0–1, 0.5–1.5, and 1.5–2.5 R_e.
STELLAR_Z	...	Flux-weighted mean redshift of the stellar component within a $2\farcs5$ aperture at the galaxy center.
STELLAR_VEL_LO	km s^{-1}	Stellar velocity at 2.5% growth of all valid spaxels.
STELLAR_VEL_HI	km s^{-1}	Stellar velocity at 97.5% growth of all valid spaxels.
STELLAR_VEL_LO_CLIP	km s^{-1}	Stellar velocity at 2.5% growth after iteratively clipping 3σ outliers.
STELLAR_VEL_HI_CLIP	km s^{-1}	Stellar velocity at 97.5% growth after iteratively clipping 3σ outliers.
STELLAR_SIGMA_1RE	km s^{-1}	Flux-weighted mean stellar velocity dispersion of all spaxels within 1 R_e.
STELLAR_CONT_RCHI2_1RE	...	Median χ^2_ν of the stellar-continuum fit within 1 R_e.
HA_Z	...	Flux-weighted mean redshift of the Hα line within a $2\farcs5$ aperture at the galaxy center.
HA_GVEL_LO	km s^{-1}	Gaussian-fitted velocity of the Hα line at 2.5% growth of all valid spaxels.
HA_GVEL_HI	km s^{-1}	Gaussian-fitted velocity of the Hα line at 97.5% growth of all valid spaxels.
HA_GVEL_LO_CLIP	km s^{-1}	Gaussian-fitted velocity of the Hα line at 2.5% growth after iteratively clipping 3σ outliers.
HA_GVEL_HI_CLIP	km s^{-1}	Gaussian-fitted velocity of the Hα line at 97.5% growth after iteratively clipping 3σ outliers.
HA_GSIGMA_1RE	km s^{-1}	Flux-weighted Hα velocity dispersion (from Gaussian fit) of all spaxels within 1 R_e.
HA_GSIGMA_HI	km s^{-1}	Hα velocity dispersion (from Gaussian fit) at 97.5% growth of all valid spaxels.
HA_GSIGMA_HI_CLIP	km s^{-1}	Hα velocity dispersion (from Gaussian fit) at 97.5% growth after iteratively clipping 3σ outliers.
EMLINE_SFLUX_CEN	10^{-17} erg s^{-1} cm^{-2}	Summed emission-line flux integrated within a $2\farcs5$ aperture at the galaxy center.
EMLINE_SFLUX_1RE	10^{-17} erg s^{-1} cm^{-2}	Summed emission-line flux integrated within 1 effective-radius aperture at the galaxy.
EMLINE_SFLUX_TOT	10^{-17} erg s^{-1} cm^{-2}	Total integrated flux of each summed emission measurement within the full MaNGA field-of-view.
EMLINE_SSB_1RE	10^{-17} erg s^{-1} cm^{-2}/spaxel^{-1}	Mean emission-line surface-brightness from the summed flux measurements within 1 R_e.
EMLINE_SSB_PEAK	10^{-17} erg s^{-1} cm^{-2}/spaxel^{-1}	Peak summed-flux emission-line surface brightness.
EMLINE_SEW_1RE	Å	Mean emission-line equivalent width from the summed flux measurements within 1 R_e.
EMLINE_SEW_PEAK	Å	Peak emission-line equivalent width from the summed flux measurements.
EMLINE_GFLUX_CEN	10^{-17} erg s^{-1} cm^{-2}	Gaussian-fitted emission-line flux integrated within a $2\farcs5$ aperture at the galaxy center.
EMLINE_GFLUX_1RE	10^{-17} erg s^{-1} cm^{-2}	Gaussian-fitted emission-line flux integrated within 1 effective-radius aperture at the galaxy.
EMLINE_GFLUX_TOT	10^{-17} erg s^{-1} cm^{-2}	Total integrated flux of the Gaussian fit to each emission line within the full MaNGA field-of-view.
EMLINE_GSB_1RE	10^{-17} erg s^{-1} cm^{-2}/spaxel^{-1}	Mean emission-line surface-brightness from the Gaussian-fitted flux measurements within 1 R_e.
EMLINE_GSB_PEAK	10^{-17} erg s^{-1} cm^{-2}/spaxel^{-1}	Peak Gaussian-fitted emission-line surface brightness.
EMLINE_GEW_1RE	Å	Mean emission-line equivalent width from the Gaussian-fitted flux measurements within 1 R_e.
EMLINE_GEW_PEAK	Å	Peak emission-line equivalent width from the Gaussian-fitted flux measurements.
SPECINDEX_LO	Å, mag	Spectral index at 2.5% growth of all valid spaxels.
SPECINDEX_HI	Å, mag	Spectral index at 97.5% growth of all valid spaxels.
SPECINDEX_LO_CLIP	Å, mag	Spectral index at 2.5% growth after iteratively clipping 3σ outliers.
SPECINDEX_HI_CLIP	Å, mag	Spectral index at 97.5% growth after iteratively clipping 3σ outliers.
SPECINDEX_1RE	Å, mag	Median spectral index within 1 effective radius.
SFR_1RE	$h^{-2}\,\mathcal{M}_\odot$/yr	Simple estimate of the star-formation rate within 1 effective radius based on the Gaussian-fitted Hα flux; logSFR = $\log L_{\mathrm{H}\alpha} - 41.27$ (Kroupa IMF; Hao et al. 2011; Murphy et al. 2011; Kennicutt & Evans 2012), where $\log L_{\mathrm{H}\alpha} = 4\pi F_{\mathrm{H}\alpha,1R_e} D_L^2$ and "*no*" attenuation correction has been applied.
SFR_TOT	$h^{-2}\,\mathcal{M}_\odot$/yr	Simple estimate of the star-formation rate within the IFU field-of-view based on the Gaussian-fitted Hα flux; logSFR = $\log L_{\mathrm{H}\alpha} - 41.27$ (Kroupa IMF; Hao et al. 2011; Murphy et al. 2011; Kennicutt & Evans 2012), where $\log L_{\mathrm{H}\alpha} = 4\pi F_{\mathrm{H}\alpha,1R_e} D_L^2$ and "*no*" attenuation correction has been applied.

Note.
[a] Calculated assuming $h = 1$, $\Omega_{\mathcal{M}} = 0.3$, and $\Omega_\Lambda = 0.7$.

ORCID iDs

Kyle B. Westfall https://orcid.org/0000-0003-1809-6920
Michele Cappellari https://orcid.org/0000-0002-1283-8420
Matthew A. Bershady https://orcid.org/0000-0002-3131-4374
Kevin Bundy https://orcid.org/0000-0001-9742-3138
Francesco Belfiore https://orcid.org/0000-0002-2545-5752
David R. Law https://orcid.org/0000-0002-9402-186X
Renbin Yan https://orcid.org/0000-0003-1025-1711
Brett H. Andrews https://orcid.org/0000-0001-8085-5890
Joel R. Brownstein https://orcid.org/0000-0002-8725-1069
Brian Cherinka https://orcid.org/0000-0002-4289-7923
Lodovico Coccato https://orcid.org/0000-0001-7817-6995
Niv Drory https://orcid.org/0000-0002-7339-3170
Claudia Maraston https://orcid.org/0000-0001-7711-3677
José R. Sánchez-Gallego https://orcid.org/0000-0003-2486-3858
Daniel Thomas https://orcid.org/0000-0002-6325-5671
Anne-Marie Weijmans https://orcid.org/0000-0002-5908-6852
Jorge Barrera-Ballesteros https://orcid.org/0000-0003-2405-7258
Niu Li https://orcid.org/0000-0003-2778-002X
Karen Masters https://orcid.org/0000-0003-0846-9578
Héctor Javier Ibarra Medel https://orcid.org/0000-0002-9790-6313
Sebastián F. Sánchez https://orcid.org/0000-0001-6444-9307
Meng Yang https://orcid.org/0000-0002-1749-1892
Shuang Zhou https://orcid.org/0000-0002-8999-6814

References

Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, ApJS, 235, 42
Aguado, D. S., Ahumada, R., Almeida, A., et al. 2019, ApJS, 240, 23
Albareti, F. D., Allende Prieto, C., Almeida, A., et al. 2017, ApJS, 233, 25
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Balogh, M. L., Morris, S. L., Yee, H. K. C., Carlberg, R. G., & Ellingson, E. 1999, ApJ, 527, 54
Beifiori, A., Maraston, C., Thomas, D., & Johansson, J. 2011, A&A, 531, A109
Belfiore, F., Westfall, K. B., & Schaefer, A. 2019, AJ, 158, 160
Bender, R. 1990, A&A, 229, 441
Blanc, G. A., Weinzirl, T., Song, M., et al. 2013, AJ, 145, 138
Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. 2017, AJ, 154, 28
Blanton, M. R., Kazin, E., Muna, D., Weaver, B. A., & Price-Whelan, A. 2011, AJ, 142, 31
Blanton, M. R., & Roweis, S. 2007, AJ, 133, 734
Bruzual, A. G. 1983, ApJ, 273, 105
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, ApJ, 798, 7
Cappellari, M. 2016, ARA&A, 54, 597
Cappellari, M. 2017, MNRAS, 466, 798
Cappellari, M., & Copin, Y. 2003, MNRAS, 342, 345
Cappellari, M., & Emsellem, E. 2004, PASP, 116, 138
Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, MNRAS, 413, 813
Cenarro, A. J., Cardiel, N., Gorgas, J., et al. 2001, MNRAS, 326, 959
Chen, Y.-M., Kauffmann, G., Tremonti, C. A., et al. 2012, MNRAS, 421, 314
Cherinka, B., Andrews, B. H., Sánchez-Gallego, J., et al. 2019, AJ, 158, 74
Conroy, C., & van Dokkum, P. 2012, ApJ, 747, 69
Croom, S. M., Lawrence, J. S., Bland-Hawthorn, J., et al. 2012, MNRAS, 421, 872
Drory, N., MacDonald, N., Bershady, M. A., et al. 2015, AJ, 149, 77
Emsellem, E., Cappellari, M., Peletier, R. F., et al. 2004, MNRAS, 352, 721
Faber, S. M., & Jackson, R. E. 1976, ApJ, 204, 668
Falcón-Barroso, J., Lyubenova, M., van de Ven, G., et al. 2017, A&A, 597, A48
Falcón-Barroso, J., Sánchez-Blázquez, P., Vazdekis, A., et al. 2011, A&A, 532, A95
Goddard, D., Thomas, D., Maraston, C., et al. 2017, MNRAS, 466, 4731
Goddard, D. S. 2018, PhD thesis, The Univ. Portsmouth
Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, AJ, 131, 2332
Hao, C.-N., Kennicutt, R. C., Johnson, B. D., et al. 2011, ApJ, 741, 124
Hogg, D. W. 1999, arXiv:astro-ph/9905116
Hunter, J. D. 2007, CSE, 9, 90
Husemann, B., Jahnke, K., Sánchez, S. F., et al. 2013, A&A, 549, A87
Jain, A. K., Murty, M. N., & Flynn, P. J. 1999, ACM Computing Surveys (CSUR), 31, 264
Johnson, S. C. 1967, Psychometrika, 32, 241
Johnston, E. J., Merrifield, M. R., Aragón-Salamanca, A., & Cappellari, M. 2013, MNRAS, 428, 1296
Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open Source Scientific Tools for Python, version 1.0.0., http://www.scipy.org/
Jorgensen, I., Franx, M., & Kjaergaard, P. 1995, MNRAS, 276, 1341
Kennicutt, R. C., & Evans, N. J. 2012, ARA&A, 50, 531
La Barbera, F., Ferreras, I., Vazdekis, A., et al. 2013, MNRAS, 433, 3017
Law, D. R., Cherinka, B., Yan, R., et al. 2016, AJ, 152, 83
Law, D. R., Yan, R., Bershady, M. A., et al. 2015, AJ, 150, 19
Lee, D. D., & Seung, H. S. 1999, Natur, 401, 788
Li, H., Mao, S., Cappellari, M., et al. 2018, MNRAS, 476, 1765
Liu, D., Blanton, M. R., & Law, D. R. 2019, arXiv:1906.06369
MacKay, D. J. C. 2003, Information Theory, Inference, and Learning Algorithms (Cambridge: Cambridge Univ. Press)
Maraston, C., & Strömbäck, G. 2011, MNRAS, 418, 2785
Markwardt, C. B. 2009, in ASP Conf. Ser. 411, Astronomical Data Analysis Software and Systems XVIII, ed. D. A. Bohlender, D. Durand, & P. Dowler (San Francisco, CA: ASP), 251
Mitzkus, M., Cappellari, M., & Walcher, C. J. 2017, MNRAS, 464, 4789
Müllner, D. 2011, arXiv:1109.2378
Murphy, E. J., Condon, J. J., Schinnerer, E., et al. 2011, ApJ, 737, 67
Nocedal, J., & Wright, S. 2006, Numerical Optimization (New York: Springer)
O'Donnell, J. E. 1994, ApJ, 422, 158
Oliphant, T. E. 2007, CSE, 9, 10
Parikh, T., Thomas, D., Maraston, C., et al. 2019, MNRAS, 483, 3420
Penny, S. J., Masters, K. L., Weijmans, A.-M., et al. 2016, MNRAS, 462, 3955
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2007, Numerical Recipes: The Art of Scientific Computing (3rd ed.; New York: Cambridge Univ. Press)
Rix, H.-W., & White, S. D. M. 1992, MNRAS, 254, 389
Ryś, A., Falcón-Barroso, J., & van de Ven, G. 2013, MNRAS, 428, 2980
Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., et al. 2012, A&A, 538, A8
Sánchez, S. F., Pérez, E., Sánchez-Blázquez, P., et al. 2016a, RMxAA, 52, 21
Sánchez, S. F., Pérez, E., Sánchez-Blázquez, P., et al. 2016b, RMxAA, 52, 171
Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., et al. 2006, MNRAS, 371, 703
Sarzi, M., Falcón-Barroso, J., Davies, R. L., et al. 2006, MNRAS, 366, 1151
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Scott, N., van de Sande, J., Croom, S. M., et al. 2018, MNRAS, 481, 2299
Serven, J., Worthey, G., & Briley, M. M. 2005, ApJ, 627, 754
Shepard, D. 1968, in Proc. 1968 23rd ACM National Conf., ed. R. B. Blue, Sr. & A. M. Rosenberg (New York: ACM), 517
Shetty, S., & Cappellari, M. 2015, MNRAS, 454, 1332
Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, AJ, 146, 32
Spiniello, C., Trager, S., Koopmans, L. V. E., & Conroy, C. 2014, MNRAS, 438, 1483
Spiniello, C., Trager, S. C., Koopmans, L. V. E., & Chen, Y. P. 2012, ApJL, 753, L32
Talbot, M. S., Brownstein, J. R., Bolton, A. S., et al. 2018, MNRAS, 477, 195
Thomas, D., Maraston, C., & Johansson, J. 2011, MNRAS, 412, 2183
Trager, S. C., Worthey, G., Faber, S. M., Burstein, D., & González, J. J. 1998, ApJS, 116, 1
Tully, R. B., & Fisher, J. R. 1977, A&A, 54, 661
van de Sande, J., Bland-Hawthorn, J., Fogarty, L. M. R., et al. 2017, ApJ, 835, 104
van der Marel, R. P., & Franx, M. 1993, ApJ, 407, 525
van der Marel, R. P., Rix, H. W., Carter, D., et al. 1994, MNRAS, 268, 521
Wake, D. A., Bundy, K., Diamond-Stanic, A. M., et al. 2017, AJ, 154, 86
Westfall, K. B., Bershady, M. A., & Verheijen, M. A. W. 2011, ApJS, 193, 21
Wilkinson, D. M., Maraston, C., Goddard, D., Thomas, D., & Parikh, T. 2017, MNRAS, 472, 4297
Willick, J. A., Courteau, S., Faber, S. M., et al. 1997, ApJS, 109, 333
Worthey, G. 1994, ApJS, 95, 107
Worthey, G., & Ottaviani, D. L. 1997, ApJS, 111, 377
Yan, R., Bundy, K., Law, D. R., et al. 2016a, AJ, 152, 197
Yan, R., Chen, Y., Lazarz, D., et al. 2019, ApJ, 883, 175
Yan, R., Newman, J. A., Faber, S. M., et al. 2006, ApJ, 648, 281
Yan, R., Tremonti, C., Bershady, M. A., et al. 2016b, AJ, 151, 8
York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579